

**PCT**

WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification <sup>7</sup> :</b> <b>C12N 9/00</b>	<b>A2</b>	<b>(11) International Publication Number:</b> <b>WO 00/18889</b> <b>(43) International Publication Date:</b> 6 April 2000 (06.04.00)
<b>(21) International Application Number:</b> PCT/US99/22231 <b>(22) International Filing Date:</b> 24 September 1999 (24.09.99)  <b>(30) Priority Data:</b> 60/101,939 25 September 1998 (25.09.98) US  <b>(71) Applicant:</b> CALGENE LLC [US/US]; 1920 Fifth Street, Davis, CA 95620 (US).  <b>(72) Inventors:</b> LASSNER, Michael, W.; 721 Falcon Avenue, Davis, CA 95616 (US). EMIG, Robin, A.; 901 Sara Court, Vacaville, CA 95687 (US). RUEZINSKY, Diane, M.; 849 Bourn Drive, Woodland, CA 95776 (US). VAN EENENNAAM, Alison; 856 Burr Street, Davis, CA 95616 (US).  <b>(74) Agents:</b> SCHWEDLER, Carl, J. et al.; Patent Dept. Central, Monsanto/G.D. Searle, P.O. Box 5110, Chicago, IL 60680-5110 (US).		<b>(81) Designated States:</b> CA, JP, MX, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).  <b>Published</b> <i>Without international search report and to be republished upon receipt of that report.</i>
<b>(54) Title:</b> NOVEL PLANT ACYLTRANSFERASES  <b>(57) Abstract</b>  By this invention, novel nucleic acid sequences encoding for acyltransferase related proteins are provided, wherein said acyltransferase-like protein is active in the transfer of a fatty acyl group from a fatty acyl donor to a fatty acyl acceptor. Also considered are amino acid and nucleic acid sequences obtainable from AT-like nucleic acid sequences and the use of such sequences to provide transgenic host cells capable of producing modified lipid content and composition.		

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TC	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakhstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

## NOVEL PLANT ACYLTRANSFERASES

5

### INTRODUCTION

This application claims the benefit of U.S. Provisional Application Serial No. 60/101,939 filed September 25, 1998.

10

#### Technical Field

The present invention is directed to nucleic acid and amino acid sequences and constructs, and methods related thereto.

15 

#### Background

Through the development of plant genetic engineering techniques, it is now possible to produce transgenic varieties of plant species to provide plants which have novel and desirable characteristics. For example, it is now possible to genetically engineer plants for tolerance to environmental stresses, such as resistance to pathogens and tolerance to herbicides and to improve the quality characteristics of the plant, for example improved fatty acid compositions. However, the number of useful nucleotide sequences for the engineering of such characteristics is thus far limited and the speed with which new useful nucleotide sequences for engineering new characteristics is slow.

The characterization of various acyltransferase proteins is useful for the further study of plant fatty acid synthesis systems and for the development of novel and/or alternative oils sources. Studies of plant mechanisms may provide means to further enhance, control, modify, or otherwise alter the total fatty acyl composition of triglycerides and oils. Furthermore, the elucidation of the factor(s) critical to the natural production of fatty acids in plants is desired, including the purification of such factors and the characterization of element(s) and/or cofactors which enhance the efficiency of the system. Of particular interest are the nucleic acid sequences of genes encoding proteins which may be useful for applications in genetic engineering.

## SUMMARY OF THE INVENTION

5           The present invention provides nucleic acid encoding for amino acid sequences for a class of proteins which are related to acyltransferase proteins. Such proteins are referred to herein as acyltransferase related or acyltransferase like proteins.

          By this invention, nucleic acid sequences encoding these acyltransferase related proteins may now be characterized with respect to enzyme activity. In particular,  
10   identification and isolation of nucleic acid sequences encoding for acyltransferase related proteins from *Arabidopsis*, yeast, corn, and soybean are provided.

          Thus, this invention encompasses acyltransferase related nucleic acid sequences and the corresponding amino acid sequences, and the use of these nucleic acid sequences in the preparation of oligonucleotides containing such acyltransferase related encoding sequences  
15   for analysis and recovery of plant acyltransferase related gene sequences. The acyltransferase related encoding sequence may encode a complete or partial sequence depending upon the intended use. All or a portion of the genomic sequence, or cDNA sequence, is intended.

          Of special interest are recombinant DNA constructs which provide for transcription or transcription and translation (expression) of the acyltransferase related sequences in host  
20   cells. In particular, constructs which are capable of transcription or transcription and translation in plant host cells are preferred. For some applications a reduction in sequences encoding acyltransferase related sequences may be desired. Thus, recombinant constructs may be designed having the acyltransferase related sequences in a reverse orientation for expression of an anti-sense sequence or use of co-suppression, also known as "transwitch",  
25   constructs may be useful. Such constructs may contain a variety of regulatory regions including transcriptional initiation regions obtained from genes preferentially expressed in plant seed tissue. For some uses, it may be desired to use the transcriptional and translational initiation regions of the acyltransferase related gene either with the acyltransferase related encoding sequence or to direct the transcription and translation of a heterologous sequence.

30           Also considered in this invention are the plants and seeds containing the constructs and polynucleotides of this invention.



## BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 provides the 204 amino acid conserved sequence profile identified from comparisons of glycerol-3-phosphate acyltransferase and various lysophosphatidic acid acyltransferase using PSI-BLAST.

Figure 2 provides an amino acid sequence alignment for the acyltransferase sequences. The alignment shown is of the regions of the protein extending from about 30 amino acids prior to the conserved H in the conserved sequence HXXXXD to 100 amino acids after, or downstream, of the P in the conserved PEG sequence motif of the acyltransferase-like sequences.

Figure 3 provides schematics showing the relationship of the identified acyltransferases. The relationships described are derived from an alignment of the regions of the protein extending from about 30 amino acids prior to the conserved H in the conserved sequence HXXXXD to 100 amino acids after, or downstream, of the P in the conserved PEG sequence motif of the acyltransferase-like sequences. Figure 3A provide a phylogenetic tree showing the relationship of several acyltransferases. Figure 3B provides a table showing the percent similarities and percent divergence of the novel acyltransferases and known acyltransferases using the Clustal method with PAM250 residue weight table.

## DETAILED DESCRIPTION OF THE INVENTION

In accordance with the subject invention, nucleotide sequences are provided which are capable of coding sequences of amino acids, such as, a protein, polypeptide or peptide, which are related to nucleic acid sequences encoding acyltransferase proteins, referred to herein as acyltransferase-like or acyltransferase related. The novel nucleic acid sequences find use in the preparation of constructs to direct their expression in a host cell. Furthermore, the novel nucleic acid sequences may find use in the preparation of plant expression constructs to modify the fatty acid composition of a plant cell.

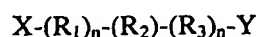
In one embodiment of the present invention, nucleic acid sequences, also referred to herein as polynucleotides, are identified from databases which are related to acyltransferases.

**Isolated proteins, Polypeptides and Polynucleotides**

A first aspect of the present invention relates to isolated acyltransferase polynucleotides. The polynucleotide sequences of the present invention include isolated  
5 polynucleotides that encode the polypeptides of the invention having a deduced amino acid sequence selected from the group of sequences set forth in the Sequence Listing and to other polynucleotide sequences closely related to such sequences and variants thereof.

The invention provides a polynucleotide sequence identical over its entire length to each coding sequence as set forth in the Sequence Listing. The invention also provides the  
10 coding sequence for the mature polypeptide or a fragment thereof, as well as the coding sequence for the mature polypeptide or a fragment thereof in a reading frame with other coding sequences, such as those encoding a leader or secretory sequence, a pre-, pro-, or prepro- protein sequence. The polynucleotide can also include non-coding sequences, including for example, but not limited to, non-coding 5' and 3' sequences, such as the  
15 transcribed, untranslated sequences, termination signals, ribosome binding sites, sequences that stabilize mRNA, introns, polyadenylation signals, and additional coding sequence that encodes additional amino acids. For example, a marker sequence can be included to facilitate the purification of the fused polypeptide. Polynucleotides of the present invention also include polynucleotides comprising a structural gene and the naturally associated sequences  
20 that control gene expression.

The invention also includes polynucleotides of the formula:



wherein, at the 5' end, X is hydrogen, and at the 3' end, Y is hydrogen or a metal,  $R_1$  and  $R_3$  are any nucleic acid residue, n is an integer between 1 and 3000, preferably between 1 and  
25 1000 and  $R_2$  is a nucleic acid sequence of the invention, particularly a nucleic acid sequence selected from the group set forth in the Sequence Listing and preferably SEQ IDNOs: 1, 3, 5, 7, 9, 10, 12, 14, 16, 18, 20, 22, and 226-233. In the formula,  $R_2$  is oriented so that its 5' end residue is at the left, bound to  $R_1$ , and its 3' end residue is at the right, bound to  $R_3$ . Any stretch of nucleic acid residues denoted by either R group, where R is greater than 1, may be  
30 either a heteropolymer or a homopolymer, preferably a heteropolymer.

The invention also relates to variants of the polynucleotides described herein that encode for variants of the polypeptides of the invention. Variants that are fragments of the polynucleotides of the invention can be used to synthesize full-length polynucleotides of the

invention. Preferred embodiments are polynucleotides encoding polypeptide variants wherein 5 to 10, 1 to 5, 1 to 3, 2, 1 or no amino acid residues of a polypeptide sequence of the invention are substituted, added or deleted, in any combination. Particularly preferred are substitutions, additions, and deletions that are silent such that they do not alter the properties or activities of the polynucleotide or polypeptide.

Nucleotide sequences encoding acyltransferases may be obtained from natural sources or be partially or wholly artificially synthesized. They may directly correspond to an acyltransferase endogenous to a natural source or contain modified amino acid sequences, such as sequences which have been mutated, truncated, increased or the like. Acyltransferases may be obtained by a variety of methods, including but not limited to, partial or homogenous purification of protein extracts, protein modeling, nucleic acid probes, antibody preparations and sequence comparisons. Typically an acyltransferase will be derived in whole or in part from a natural source. A natural source includes, but is not limited to, prokaryotic and eukaryotic sources, including, bacteria, yeasts, plants, including algae, and the like.

Of special interest are acyltransferases which are obtainable from eukaryotic sources, including those which are obtained, from plants, or from acyltransferases which are obtainable through the use of these sequences. "Obtainable" refers to those acyltransferases which have sufficiently similar sequences to that of the sequences provided herein to provide a biologically active protein of the present invention.

Further preferred embodiments of the invention that are at least 50%, 60%, or 70% identical over their entire length to a polynucleotide encoding a polypeptide of the invention, and polynucleotides that are complementary to such polynucleotides. More preferable are polynucleotides that comprise a region that is at least 80% identical over its entire length to a polynucleotide encoding a polypeptide of the invention and polynucleotides that are complementary thereto. In this regard, polynucleotides at least 90% identical over their entire length are particularly preferred, those at least 95% identical are especially preferred. Further, those with at least 97% identity are highly preferred and those with at least 98% and 99% identity are particularly highly preferred, with those at least 99% being the most highly preferred.

Preferred embodiments are polynucleotides that encode polypeptides that retain substantially the same biological function or activity as the mature polypeptides encoded by the polynucleotides set forth in the Sequence Listing.

The invention further relates to polynucleotides that hybridize to the above-described sequences. In particular, the invention relates to polynucleotides that hybridize under stringent conditions to the above-described polynucleotides. As used herein, the terms "stringent conditions" and "stringent hybridization conditions" mean that hybridization will generally occur if there is at least 95% and preferably at least 97% identity between the sequences. An example of stringent hybridization conditions is overnight incubation at 42°C in a solution comprising 50% formamide, 5x SSC (150 mM NaCl, 15 mM trisodium citrate), 50 mM sodium phosphate (pH 7.6), 5x Denhardt's solution, 10% dextran sulfate, and 20 micrograms/milliliter denatured, sheared salmon sperm DNA, followed by washing the hybridization support in 0.1x SSC at approximately 65°C. Other hybridization and wash conditions are well known and are exemplified in Sambrook, *et al.*, Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor, NY (1989), particularly Chapter 11.

The invention also provides a polynucleotide consisting essentially of a polynucleotide sequence obtainable by screening an appropriate library containing the complete gene for a polynucleotide sequence set forth in the Sequence Listing under stringent hybridization conditions with a probe having the sequence of said polynucleotide sequence or a fragment thereof; and isolating said polynucleotide sequence. Fragments useful for obtaining such a polynucleotide include, for example, probes and primers as described herein.

As discussed herein regarding polynucleotide assays of the invention, for example, polynucleotides of the invention can be used as a hybridization probe for RNA, cDNA, or genomic DNA to isolate full length cDNAs or genomic clones encoding a polypeptide and to isolate cDNA or genomic clones of other genes that have a high sequence similarity to a polynucleotide set forth in the Sequence Listing. Such probes will generally comprise at least 15 bases. Preferably such probes will have at least 30 bases and can have at least 50 bases. Particularly preferred probes will have between 30 bases and 50 bases, inclusive.

The coding region of each gene that comprises or is comprised by a polynucleotide sequence set forth in the Sequence Listing may be isolated by screening using a DNA sequence provided in the Sequence Listing to synthesize an oligonucleotide probe. A labeled oligonucleotide having a sequence complementary to that of a gene of the invention is then used to screen a library of cDNA, genomic DNA or mRNA to identify members of the library which hybridize to the probe. For example, synthetic oligonucleotides are prepared which correspond to the N-terminal sequence of the polypeptide. The partial sequences so prepared can then be used as probes to obtain acyltransferase clones from a gene library prepared from

a cell source of interest. Alternatively, where oligonucleotides of low degeneracy can be prepared from particular peptides, such probes may be used directly to screen gene libraries for gene sequences. In particular, screening of cDNA libraries in phage vectors is useful in such methods due to lower levels of background hybridization.

5           Typically, a sequence obtainable from the use of nucleic acid probes will show 60-70% sequence identity between the target acyltransferase sequence and the encoding sequence used as a probe. However, lengthy sequences with as little as 50-60% sequence identity may also be obtained. The nucleic acid probes may be a lengthy fragment of the nucleic acid sequence, or may also be a shorter, oligonucleotide probe. When longer nucleic acid  
10 fragments are employed as probes (greater than about 100 bp), one may screen at lower stringencies in order to obtain sequences from the target sample which have 20-50% deviation (i.e., 50-80% sequence homology) from the sequences used as probe. Oligonucleotide probes can be considerably shorter than the entire nucleic acid sequence encoding an acyltransferase enzyme, but should be at least about 10, preferably at least about  
15 15, and more preferably at least about 20 nucleotides. A higher degree of sequence identity is desired when shorter regions are used as opposed to longer regions. It may thus be desirable to identify regions of highly conserved amino acid sequence to design oligonucleotide probes for detecting and recovering other related genes. Shorter probes are often particularly useful for polymerase chain reactions (PCR), especially when highly conserved sequences can be  
20 identified. (*See, Gould, et al., PNAS USA (1989) 86:1934-1938*).

The skilled artisan will appreciate that, in many cases, an isolated cDNA sequence will be incomplete, in that the region coding for the polypeptide is truncated with respect to the 5' terminus of the cDNA. This is a consequence of the reverse transcriptase, an enzyme with low 'processivity' (a measure of the ability of the enzyme to remain attached to the  
25 template during the polymerization reaction) employed during the first strand cDNA synthesis.

There are several methods available and are well known to the skilled artisan to obtain full-length cDNAs, or extend short cDNAs, for example those based on the method of Rapid Amplification of cDNA Ends (RACE) (see, for example, Frohman *et al.* (1988) *Proc. Natl. Acad. Sci. USA* 85:8998-9002). Recent modifications of the technique, exemplified by the  
30 Marathon™ technology (Clontech Laboratories, Inc.) for example, have significantly simplified obtaining full-length cDNA sequences.

Another aspect of the present invention relates to isolated acyltransferase polypeptides. Such polypeptides include isolated polypeptides set forth in the Sequence Listing, as well as polypeptides and fragments thereof, particularly those polypeptides which exhibit acyltransferase activity and also those polypeptides which have at least 50%, 60% or 70% identity, preferably at least 80% identity, more preferably at least 90% identity, and most preferably at least 95% identity to a polypeptide sequence selected from the group of sequences set forth in the Sequence Listing, and also include portions of such polypeptides, wherein such portion of the polypeptide preferably includes at least 30 amino acids and more preferably includes at least 50 amino acids.

“Identity”, as is well understood in the art, is a relationship between two or more polypeptide sequences or two or more polynucleotide sequences, as determined by comparing the sequences. In the art, “identity” also means the degree of sequence relatedness between polypeptide or polynucleotide sequences, as determined by the match between strings of such sequences. “Identity” can be readily calculated by known methods including, but not limited to, those described in *Computational Molecular Biology*, Lesk, A.M., ed., Oxford University Press, New York (1988); *Biocomputing: Informatics and Genome Projects*, Smith, D.W., ed., Academic Press, New York, 1993; *Computer Analysis of Sequence Data, Part I*, Griffin, A.M. and Griffin, H.G., eds., Humana Press, New Jersey (1994); *Sequence Analysis in Molecular Biology*, von Heinje, G., Academic Press (1987); *Sequence Analysis Primer*, Gribskov, M. and Devereux, J., eds., Stockton Press, New York (1991); and Carillo, H., and Lipman, D., *SIAM J Applied Math*, 48:1073 (1988). Methods to determine identity are designed to give the largest match between the sequences tested. Moreover, methods to determine identity are codified in publicly available programs. Computer programs which can be used to determine identity between two sequences include, but are not limited to, GCG (Devereux, J., et al., *Nucleic Acids Research* 12(1):387 (1984); suite of five BLAST programs, three designed for nucleotide sequences queries (BLASTN, BLASTX, and TBLASTX) and two designed for protein sequence queries (BLASTP and TBLASTN) (Coulson, *Trends in Biotechnology*, 12: 76-80 (1994); Birren, et al., *Genome Analysis*, 1: 543-559 (1997)). The BLAST X program is publicly available from NCBI and other sources (BLAST Manual, Altschul, S., et al., NCBI NLM NIH, Bethesda, MD 20894; Altschul, S., et al., *J. Mol. Biol.*, 215:403-410 (1990)). The well known Smith Waterman algorithm can also be used to determine identity.

Parameters for polypeptide sequence comparison typically include the following:

Algorithm: Needleman and Wunsch, *J. Mol. Biol.* 48:443-453 (1970)

Comparison matrix: BLOSSUM62 from Hentikoff and Hentikoff, *Proc. Natl. Acad. Sci USA* 89:10915-10919 (1992)

5           Gap Penalty: 12

          Gap Length Penalty: 4

A program which can be used with these parameters is publicly available as the "gap" program from Genetics Computer Group, Madison Wisconsin. The above parameters along with no penalty for end gap are the default parameters for peptide comparisons.

10           Parameters for polynucleotide sequence comparison include the following:

Algorithm: Needleman and Wunsch, *J. Mol. Biol.* 48:443-453 (1970)

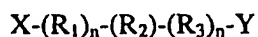
Comparison matrix: matches = +10; mismatches = 0

Gap Penalty: 50

Gap Length Penalty: 3

15           A program which can be used with these parameters is publicly available as the "gap" program from Genetics Computer Group, Madison Wisconsin. The above parameters are the default parameters for nucleic acid comparisons.

The invention also includes polypeptides of the formula:



20   wherein, at the amino terminus, X is hydrogen, and at the carboxyl terminus, Y is hydrogen or a metal,  $R_1$  and  $R_3$  are any amino acid residue,  $n$  is an integer between 1 and 1000, and  $R_2$  is an amino acid sequence of the invention, particularly an amino acid sequence selected from the group set forth in the Sequence Listing and preferably SEQ IDNOs: 2, 4, 6, 8, 11, 13, 15, 17, 19, 21, 23, and 218-225. In the formula,  $R_2$  is oriented so that its amino terminal residue  
25   is at the left, bound to  $R_1$ , and its carboxy terminal residue is at the right, bound to  $R_3$ . Any stretch of amino acid residues denoted by either R group, where R is greater than 1, may be either a heteropolymer or a homopolymer, preferably a heteropolymer.

Polypeptides of the present invention include isolated polypeptides encoded by a polynucleotide comprising a sequence selected from the group of a sequence contained in  
30   SEQ ID NOs: 1, 3, 5, 7, 9, 10, 12, 14, 16, 18, 20, 22, and 226-233.

The polypeptides of the present invention can be mature protein or can be part of a fusion protein.

Fragments and variants of the polypeptides are also considered to be a part of the invention. A fragment is a variant polypeptide which has an amino acid sequence that is entirely the same as part but not all of the amino acid sequence of the previously described polypeptides. The fragments can be "free-standing" or comprised within a larger polypeptide of which the fragment forms a part or a region, most preferably as a single continuous region. Preferred fragments are biologically active fragments which are those fragments that mediate activities of the polypeptides of the invention, including those with similar activity or improved activity or with a decreased activity. Also included are those fragments that antigenic or immunogenic in an animal, particularly a human.

Variants of the polypeptide also include polypeptides that vary from the sequences set forth in the Sequence Listing by conservative amino acid substitutions, substitution of a residue by another with like characteristics. In general, such substitutions are among Ala, Val, Leu and Ile; between Ser and Thr; between Asp and Glu; between Asn and Gln; between Lys and Arg; or between Phe and Tyr. Particularly preferred are variants in which 5 to 10; 1 to 5; 1 to 3 or one amino acid(s) are substituted, deleted, or added, in any combination.

Variants that are fragments of the polypeptides of the invention can be used to produce the corresponding full length polypeptide by peptide synthesis. Therefore, these variants can be used as intermediates for producing the full-length polypeptides of the invention.

The polynucleotides and polypeptides of the invention can be used, for example, in the transformation of various host cells, as further discussed herein.

The invention also provides polynucleotides that encode a polypeptide that is a mature protein plus additional amino or carboxyl-terminal amino acids, or amino acids within the mature polypeptide (for example, when the mature form of the protein has more than one polypeptide chain). Such sequences can, for example, play a role in the processing of a protein from a precursor to a mature form, allow protein transport, shorten or lengthen protein half-life, or facilitate manipulation of the protein in assays or production. It is contemplated that cellular enzymes can be used to remove any additional amino acids from the mature protein.

A precursor protein, having the mature form of the polypeptide fused to one or more prosequences may be an inactive form of the polypeptide. The inactive precursors generally are activated when the prosequences are removed. Some or all of the prosequences may be removed prior to activation. Such precursor protein are generally called proproteins.



The polynucleotide and polypeptide sequences can also be used to identify additional sequences which are homologous to the sequences of the present invention. The most preferable and convenient method is to store the sequence in a computer readable medium, for example, floppy disk, CD ROM, hard disk drives, external disk drives and DVD, and then to use the stored sequence to search a sequence database with well known searching tools.

Examples of public databases include the DNA Database of Japan

(DDBJ)(<http://www.ddbj.nig.ac.jp/>); Genebank

(<http://www.ncbi.nlm.nih.gov/web/Genbank/Index.html>); and the European Molecular Biology Laboratory Nucleic Acid Sequence Database (EMBL)

([http://www.ebi.ac.uk/ebi\\_docs/embl\\_db.html](http://www.ebi.ac.uk/ebi_docs/embl_db.html)). A number of different search algorithms are available to the skilled artisan, one example of which are the suite of programs referred to as BLAST programs. There are five implementations of BLAST, three designed for nucleotide sequences queries (BLASTN, BLASTX, and TBLASTX) and two designed for protein sequence queries (BLASTP and TBLASTN) (Coulson, *Trends in Biotechnology*, 12: 76-80 (1994); Birren, *et al.*, *Genome Analysis*, 1: 543-559 (1997)). Additional programs are available in the art for the analysis of identified sequences, such as sequence alignment programs, programs for the identification of more distantly related sequences, and the like, and are well known to the skilled artisan.

## 20 Plant Constructs and Methods of Use

Of interest in the present invention, is the use of the nucleotide sequences, or polynucleotides, in recombinant DNA constructs to direct the transcription or transcription and translation (expression) of the acyltransferase sequences of the present invention in a host cell.

Of particular interest is the use of the nucleotide sequences, or polynucleotides, in recombinant DNA constructs to direct the transcription or transcription and translation (expression) of the acyltransferase sequences of the present invention in a host cell. The expression constructs generally comprise a promoter functional in a host cell operably linked to a nucleic acid sequence encoding an acyltransferase of the present invention and a transcriptional termination region functional in a host cell.

By "host cell" is meant a cell which contains a vector and supports the replication, and/or transcription or transcription and translation (expression) of the expression construct.

Host cells for use in the present invention can be prokaryotic cells, such as *E. coli*, or eukaryotic cells such as yeast, plant, insect, amphibian, or mammalian cells. Preferably, host cells are monocotyledenous or dicotyledenous plant cells.

Of particular interest in the present invention is the use of the polynucleotides of the present invention for the preparation of constructs to direct the transcription or transcription and translation of the nucleotide sequences encoding an acyltransferase in a host plant cell. Plant expression constructs generally comprise a promoter functional in a plant host cell operably linked to a nucleic acid sequence of the present and a transcriptional termination region functional in a host plant cell.

Those skilled in the art will recognize that there are a number of promoters which are functional in plant cells, and have been described in the literature. Chloroplast and plastid specific promoters, chloroplast or plastid functional promoters, and chloroplast or plastid operable promoters are also envisioned.

One set of promoters are constitutive promoters such as the CaMV35S or FMV35S promoters that yield high levels of expression in most plant organs. Enhanced or duplicated versions of the CaMV35S and FMV35S promoters are useful in the practice of this invention (Odell, *et al.* (1985) *Nature* 313:810-812; Rogers, U.S. Patent Number 5,378, 619). In addition, it may also be preferred to bring about expression of the protein of interest in specific tissues of the plant, such as leaf, stem, root, tuber, seed, fruit, etc., and the promoter chosen should have the desired tissue and developmental specificity.

Of particular interest is the expression of the nucleic acid sequences of the present invention from transcription initiation regions which are preferentially expressed in a plant seed tissue. Examples of such seed preferential transcription initiation sequences include those sequences derived from sequences encoding plant storage protein genes or from genes involved in fatty acid biosynthesis in oilseeds. Examples of such promoters include the 5' regulatory regions from such genes as napin (Kridl *et al.*, *Seed Sci. Res.* 1:209:219 (1991)), phaseolin, zein, soybean trypsin inhibitor, ACP, stearyl-ACP desaturase, soybean  $\alpha'$  subunit of  $\beta$ -conglycinin (soy 7s, (Chen *et al.*, *Proc. Natl. Acad. Sci.*, 83:8560-8564 (1986))) and oleosin.

It may be advantageous to direct the localization of proteins conferring acyltransferase to a particular subcellular compartment, for example, to the mitochondrion, endoplasmic reticulum, vacuoles, chloroplast or other plastidic compartment. For example, where the genes of interest of the present invention will be targeted to plastids, such as chloroplasts, for

expression, the constructs will also employ the use of sequences to direct the gene to the plastid. Such sequences are referred to herein as chloroplast transit peptides (CTP) or plastid transit peptides (PTP). In this manner, where the gene of interest is not directly inserted into the plastid, the expression construct will additionally contain a gene encoding a transit peptide to direct the gene of interest to the plastid. The chloroplast transit peptides may be derived from the gene of interest, or may be derived from a heterologous sequence having a CTP. Such transit peptides are known in the art. See, for example, Von Heijne *et al.* (1991) *Plant Mol. Biol. Rep.* 9:104-126; Clark *et al.* (1989) *J. Biol. Chem.* 264:17544-17550; della-Cioppa *et al.* (1987) *Plant Physiol.* 84:965-968; Romer *et al.* (1993) *Biochem. Biophys. Res Commun.* 196:1414-1421; and, Shah *et al.* (1986) *Science* 233:478-481. Additional transit peptides for the translocation of the protein to the endoplasmic reticulum (ER), or vacuole may also find use in the constructs of the present invention.

Depending upon the intended use, the constructs may contain the nucleic acid sequence which encodes the entire acyltransferase protein, or a portion thereof. For example, where antisense inhibition of a given acyltransferase protein is desired, the entire sequence is not required. Furthermore, where acyltransferase sequences used in constructs are intended for use as probes, it may be advantageous to prepare constructs containing only a particular portion of a acyltransferase encoding sequence, for example a sequence which is discovered to encode a highly conserved acyltransferase region.

The skilled artisan will recognize that there are various methods for the inhibition of expression of endogenous sequences in a host cell. Such methods include, but are not limited to antisense suppression (Smith, *et al.* (1988) *Nature* 334:724-726), co-suppression (Napoli, *et al.* (1989) *Plant Cell* 2:279-289), ribozymes (PCT Publication WO 97/10328), and combinations of sense and antisense, such as those described by Waterhouse, *et al.* (1998) *Proc. Natl. Acad. Sci. USA* 95:13959-13964. Methods for the suppression of endogenous sequences in a host cell typically employ the transcription or transcription and translation of at least a portion of the sequence to be suppressed. Such sequences may be homologous to coding as well as non-coding regions of the endogenous sequence.

Regulatory transcript termination regions may be provided in plant expression constructs of this invention as well. Transcript termination regions may be provided by the DNA sequence encoding the acyltransferase or a convenient transcription termination region derived from a different gene source, for example, the transcript termination region which is naturally associated with the transcript initiation region. The skilled artisan will recognize

that any convenient transcript termination region which is capable of terminating transcription in a plant cell may be employed in the constructs of the present invention.

Alternatively, constructs may be prepared to direct the expression of the acyltransferase sequences directly from the host plant cell plastid. Such constructs and methods are known in the art and are generally described, for example, in Svab, *et al.* (1990) *Proc. Natl. Acad. Sci. USA* 87:8526-8530 and Svab and Maliga (1993) *Proc. Natl. Acad. Sci. USA* 90:913-917 and in U.S. Patent Number 5,693,507.

A plant cell, tissue, organ, or plant into which the recombinant DNA constructs containing the expression constructs have been introduced is considered transformed, transfected, or transgenic. A transgenic or transformed cell or plant also includes progeny of the cell or plant and progeny produced from a breeding program employing such a transgenic plant as a parent in a cross and exhibiting an altered genotype resulting from the presence of an introduced acyltransferase nucleic acid sequence.

The term "introduced" in the context of inserting a nucleic acid sequence into a cell, means "transfection", or "transformation" or "transduction" and includes reference to the incorporation of a nucleic acid sequence into a eukaryotic or prokaryotic cell where the nucleic acid sequence may be incorporated into the genome of the cell (for example, chromosome, plasmid, plastid, or mitochondrial DNA), converted into an autonomous replicon, or transiently expressed (for example, transfected mRNA).

Plant expression or transcription constructs having an acyltransferase as the DNA sequence of interest for increased or decreased expression thereof may be employed with a wide variety of plant life, particularly, plant life involved in the production of vegetable oils for edible and industrial uses. Plants of interest in the present invention include monocotyledenous and dicotyledenous plants. Most especially preferred are temperate oilseed crops. Plants of interest include, but are not limited to, rapeseed (Canola and High Erucic Acid varieties), sunflower, safflower, cotton, soybean, peanut, coconut and oil palms, and corn. Depending on the method for introducing the recombinant constructs into the host cell, other DNA sequences may be required. Importantly, this invention is applicable to dicotyledons and monocotyledons species alike and will be readily applicable to new and/or improved transformation and regulation techniques.

As used herein, the term "plant" includes reference to whole plants, plant organs (for example, leaves, stems, roots, etc.), seeds, and plant cells and progeny of same. Plant cell, as used herein includes, without limitation, seeds suspension cultures, embryos, meristematic

regions, callus tissue, leaves roots shoots, gametophytes, sporophytes, pollen, and microspores. The class of plants which can be used in the methods of the present invention is generally as broad as the class of higher plants amenable to transformation techniques, including both monocotyledenous and dicotyledenous plants. Particularly preferred plants of interest include, but are not limited to, rapeseed (Canola and High Erucic Acid varieties), sunflower, safflower, cotton, soybean, peanut, coconut and oil palms, and corn. Most especially preferred plants include *Brassica*, soybean, and corn.

As used herein, "transgenic plant" includes reference to a plant which comprises within its genome a heterologous polynucleotide. Generally, the heterologous polynucleotide is stably integrated within the genome such that the polynucleotide is passed on to successive generations. The heterologous polynucleotide may be integrated into the genome alone or as part of a recombinant expression cassette. "Transgenic" is used herein to include any cell, cell line, callus, tissue, plant part or plant, the genotype of which has been altered by the presence of heterologous nucleic acid including those transgenics initially so altered as well as those created by sexual crosses or asexual propagation from the initial transgenic.

Thus a plant having within its cells a heterologous polynucleotide is referred to herein as a transgenic plant. The heterologous polynucleotide can be either stably integrated into the genome, or can be extra-chromosomal. Preferably, the polynucleotide of the present invention is stably integrated into the genome such that the polynucleotide is passed on to successive generations. The polynucleotide is integrated into the genome alone or as part of a recombinant expression cassette. "Transgenic" is used herein to include any cell, cell line, callus, tissue, plant part or plant, the genotype of which has been altered by the presence of heterologous nucleic acids including those transgenics initially so altered as well as those created by sexual crosses or asexual reproduction of the initial transgenics.

As used herein, "heterologous" in reference to a nucleic acid is a nucleic acid that originates from a foreign species, or, if from the same species, is substantially modified from its native form in composition and/or genomic locus by deliberate human intervention. For example, a promoter operably linked to a heterologous structural gene is from a species different from that from which the structural gene was derived, or, if from the same species, one or both are substantially modified from their original form. A heterologous protein may originate from a foreign species, or, if from the same species, is substantially modified from its original form by deliberate human intervention.

As used herein, a "recombinant expression cassette" is a nucleic acid construct, generated recombinantly or synthetically, with a series of specified nucleic acid elements which permit transcription of a particular nucleic acid in a target cell. The recombinant expression cassette can be incorporated into a plasmid, chromosome, mitochondrial DNA, plastid DNA, virus, or nucleic acid fragment. Typically, the recombinant expression cassette portion of an expression vector includes, among other sequences, a nucleic acid sequence to be transcribed and a promoter.

It is contemplated that the gene sequences may be synthesized, either completely or in part, especially where it is desirable to provide plant-preferred sequences. Thus, all or a portion of the desired structural gene (that portion of the gene which encodes the acyltransferase protein) may be synthesized using codons preferred by a selected host. Host-preferred codons may be determined, for example, from the codons used most frequently in the proteins expressed in a desired host species.

One skilled in the art will readily recognize that antibody preparations, nucleic acid probes (DNA and RNA) and the like may be prepared and used to screen and recover "homologous" or "related" acyltransferase from a variety of plant sources. Homologous sequences are found when there is an identity of sequence, which may be determined upon comparison of sequence information, nucleic acid or amino acid, or through hybridization reactions between a known acyltransferase and a candidate source. Conservative changes, such as Glu/Asp, Val/Ile, Ser/Thr, Arg/Lys and Gln/Asn may also be considered in determining sequence homology. Amino acid sequences are considered homologous by as little as 25% sequence identity between the two complete mature proteins. (See generally, Doolittle, R.F., *OF URFS and ORFS* (University Science Books, CA, 1986.)

Thus, other acyltransferase sequences can be obtained from the specific exemplified sequences provided herein. Furthermore, it will be apparent that one can obtain natural and synthetic sequences, including modified amino acid sequences and starting materials for synthetic-protein modeling from the exemplified sequences and from acyltransferases which are obtained through the use of such exemplified sequences. Modified amino acid sequences include sequences which have been mutated, truncated, increased and the like, whether such sequences were partially or wholly synthesized. Sequences which are actually purified from plant preparations or are identical or encode identical proteins thereto, regardless of the method used to obtain the protein or sequence, are equally considered naturally derived.

For immunological screening, antibodies to the acyltransferase protein can be prepared by injecting rabbits or mice with the purified protein or portion thereof, such methods of preparing antibodies being well known to those in the art. Either monoclonal or polyclonal antibodies can be produced, although typically polyclonal antibodies are more useful for gene isolation. Western analysis may be conducted to determine that a related protein is present in a crude extract of the desired plant species, as determined by cross-reaction with the antibodies to the acyltransferase protein. When cross-reactivity is observed, genes encoding the related proteins are isolated by screening expression libraries representing the desired plant species. Expression libraries can be constructed in a variety of commercially available vectors, including lambda gt11, as described in Sambrook, *et al.* (*Molecular Cloning: A Laboratory Manual*, Second Edition (1989) Cold Spring Harbor Laboratory, Cold Spring Harbor, New York).

The nucleic acid sequences associated with acyltransferase proteins will find many uses. For example, recombinant constructs can be prepared which can be used as probes, or which will provide for expression of the acyltransferase protein in host cells to produce a ready source of the enzyme and/or to modify the composition of triglycerides found therein. Other useful applications may be found when the host cell is a plant host cell, either *in vitro* or *in vivo*.

The modification of fatty acid compositions may also affect the fluidity of plant membranes. Different lipid concentrations have been observed in cold-hardened plants, for example. By this invention, one may be capable of introducing traits which will lend to chill tolerance. Constitutive or temperature inducible transcription initiation regulatory control regions may have special applications for such uses.

As discussed above, nucleic acid sequence encoding an acyltransferase of this invention may include genomic, cDNA or mRNA sequence. By "encoding" is meant that the sequence corresponds to a particular amino acid sequence either in a sense or anti-sense orientation. By "extrachromosomal" is meant that the sequence is outside of the plant genome of which it is naturally associated. By "recombinant" is meant that the sequence contains a genetically engineered modification through manipulation via mutagenesis, restriction enzymes, and the like.

Once the desired acyltransferase nucleic acid sequence is obtained, it may be manipulated in a variety of ways. Where the sequence involves non-coding flanking regions, the flanking regions may be subjected to resection, mutagenesis, etc. Thus, transitions,

transversions, deletions, and insertions may be performed on the naturally occurring sequence. In addition, all or part of the sequence may be synthesized. In the structural gene, one or more codons may be modified to provide for a modified amino acid sequence, or one or more codon mutations may be introduced to provide for a convenient restriction site or  
5 other purpose involved with construction or expression. The structural gene may be further modified by employing synthetic adapters, linkers to introduce one or more convenient restriction sites, or the like.

The nucleic acid or amino acid sequences encoding an acyltransferase of this invention may be combined with other non-native, or "heterologous", sequences in a variety  
10 of ways. By "heterologous" sequences is meant any sequence which is not naturally found joined to the acyltransferase, including, for example, combinations of nucleic acid sequences from the same plant which are not naturally found joined together.

The DNA sequence encoding an acyltransferase of this invention may be employed in conjunction with all or part of the gene sequences normally associated with the  
15 acyltransferase. In its component parts, a DNA sequence encoding acyltransferase is combined in a DNA construct having, in the 5' to 3' direction of transcription, a transcription initiation control region capable of promoting transcription and translation in a host cell, the DNA sequence encoding plant acyltransferase and a transcription and translation termination region.

Potential host cells include both prokaryotic cells, such as *E.coli* and eukaryotic cells  
20 such as yeast, insect, amphibian, or mammalian cells. A host cell may be unicellular or found in a multicellular differentiated or undifferentiated organism depending upon the intended use. Preferably, host cells of the present invention include plant cells, both monocotyledenous and dicotyledenous. Cells of this invention may be distinguished by  
25 having a sequence foreign to the wild-type cell present therein, for example, by having a recombinant nucleic acid construct encoding an acyltransferase therein.

The methods used for the transformation of the host plant cell are not critical to the present invention. The transformation of the plant is preferably permanent, i.e. by integration of the introduced expression constructs into the host plant genome, so that the introduced  
30 constructs are passed onto successive plant generations. The skilled artisan will recognize that a wide variety of transformation techniques exist in the art, and new techniques are continually becoming available. Any technique that is suitable for the target host plant can be employed within the scope of the present invention. For example, the constructs can be



introduced in a variety of forms including, but not limited to as a strand of DNA, in a plasmid, or in an artificial chromosome. The introduction of the constructs into the target plant cells can be accomplished by a variety of techniques, including, but not limited to calcium-phosphate-DNA co-precipitation, electroporation, microinjection, *Agrobacterium* infection, liposomes or microprojectile transformation. The skilled artisan can refer to the literature for details and select suitable techniques for use in the methods of the present invention.

Normally, included with the DNA construct will be a structural gene having the necessary regulatory regions for expression in a host and providing for selection of transformant cells. The gene may provide for resistance to a cytotoxic agent, e.g. antibiotic, heavy metal, toxin, etc., complementation providing prototrophy to an auxotrophic host, viral immunity or the like. Depending upon the number of different host species the expression construct or components thereof are introduced, one or more markers may be employed, where different conditions for selection are used for the different hosts.

Where *Agrobacterium* is used for plant cell transformation, a vector may be used which may be introduced into the *Agrobacterium* host for homologous recombination with T-DNA or the Ti- or Ri-plasmid present in the *Agrobacterium* host. The Ti- or Ri-plasmid containing the T-DNA for recombination may be armed (capable of causing gall formation) or disarmed (incapable of causing gall formation), the latter being permissible, so long as the *vir* genes are present in the transformed *Agrobacterium* host. The armed plasmid can give a mixture of normal plant cells and gall.

In some instances where *Agrobacterium* is used as the vehicle for transforming host plant cells, the expression or transcription construct bordered by the T-DNA border region(s) will be inserted into a broad host range vector capable of replication in *E. coli* and *Agrobacterium*, there being broad host range vectors described in the literature. Commonly used is pRK2 or derivatives thereof. See, for example, Ditta, *et al.*, (*Proc. Nat. Acad. Sci., U.S.A.* (1980) 77:7347-7351) and EPA 0 120 515, which are incorporated herein by reference. Alternatively, one may insert the sequences to be expressed in plant cells into a vector containing separate replication sequences, one of which stabilizes the vector in *E. coli*, and the other in *Agrobacterium*. See, for example, McBride and Summerfelt (*Plant Mol. Biol.* (1990) 14:269-276), wherein the pRiHRI (Jouanin, *et al.*, *Mol. Gen. Genet.* (1985) 201:370-374) origin of replication is utilized and provides for added stability of the plant expression vectors in host *Agrobacterium* cells.

Included with the expression construct and the T-DNA will be one or more markers, which allow for selection of transformed *Agrobacterium* and transformed plant cells. A number of markers have been developed for use with plant cells, such as resistance to chloramphenicol, kanamycin, the aminoglycoside G418, hygromycin, or the like. The particular marker employed is not essential to this invention, one or another marker being preferred depending on the particular host and the manner of construction.

For transformation of plant cells using *Agrobacterium*, explants may be combined and incubated with the transformed *Agrobacterium* for sufficient time for transformation, the bacteria killed, and the plant cells cultured in an appropriate selective medium. Once callus forms, shoot formation can be encouraged by employing the appropriate plant hormones in accordance with known methods and the shoots transferred to rooting medium for regeneration of plants. The plants may then be grown to seed and the seed used to establish repetitive generations and for isolation of vegetable oils.

There are several possible ways to obtain the plant cells of this invention which contain multiple expression constructs. Any means for producing a plant comprising a construct having a nucleic acid sequence of the present invention, and at least one other construct having another DNA sequence encoding an enzyme are encompassed by the present invention. For example, the expression construct can be used to transform a plant at the same time as the second construct either by inclusion of both expression constructs in a single transformation vector or by using separate vectors, each of which express desired genes. The second construct can be introduced into a plant which has already been transformed with the first expression construct, or alternatively, transformed plants, one having the first construct and one having the second construct, can be crossed to bring the constructs together in the same plant.

In general, acyltransferase proteins are active in the transfer of acyl groups from a donor to a variety of different substrates. For example, diacylglycerol acyltransferases add acyl groups to diacylglycerol to form triacylglycerol (TAG), or acyl:CoA:cholesterol acyltransferase uses an acyl-CoA as a donor to transfer an acyl group to a sterol to form a sterol ester. Typically, the substrates include, but are not limited to glycerides, including mono and diglycerides, sterols, stanols, phosphatides, and the like. Donors include, but are not limited to acyl-CoA and acyl-ACP molecules.

The invention now being generally described, it will be more readily understood by reference to the following examples which are included for purposes of illustration only and are not intended to limit the present invention.

5

## EXAMPLES

### Example 1: RNA Isolations

10           Total RNA from the inflorescence and developing seeds of *Arabidopsis thaliana* is isolated for use in construction of complementary (cDNA) libraries. The procedure is an adaptation of the DNA isolation protocol of Webb and Knapp (D.M. Webb and S.J. Knapp, (1990) Plant Molec. Reporter, 8, 180-185). The following description assumes the use of 1g fresh weight of tissue. Frozen seed tissue is powdered by grinding under liquid nitrogen. The powder is added to 10ml REC buffer (50mM Tris-HCl, pH 9, 0.8M NaCl, 10mM EDTA, 0.5% w/v CTAB (cetyltrimethyl-ammonium bromide)) along with 0.2g insoluble polyvinylpolypyrrolidone, and ground at room temperature. The homogenate is centrifuged for 5 minutes at 12,000 xg to pellet insoluble material. The resulting supernatant fraction is extracted with chloroform, and the top phase is recovered.

20           The RNA is then precipitated by addition of 1 volume RecP (50mM Tris-HCL pH9, 10mM EDTA and 0.5% (w/v) CTAB) and collected by brief centrifugation as before. The RNA pellet is redissolved in 0.4 ml of 1M NaCl. The RNA pellet is redissolved in water and extracted with phenol/chloroform. Sufficient 3M potassium acetate (pH 5) is added to make the mixture 0.3M in acetate, followed by addition of two volumes of ethanol to precipitate the RNA. After washing with ethanol, this final RNA precipitate is dissolved in water and stored frozen.

          Alternatively, total RNA may be obtained using TRIzol reagent (BRL-Lifetechnologies, Gaithersburg, MD) following the manufacturers protocol. The RNA precipitate is dissolved in water and stored frozen.

30

### Example 2: Identification of Acyltransferase Homology Sequences

Searches are performed on a Silicon Graphics Unix computer using additional Bioaccelerator hardware and GenWeb software supplied by Compugen Ltd. This software and hardware enables the use of the Smith-Waterman algorithm in searching DNA and protein databases using profiles as queries. The program used to query protein databases is profilesearch. This is a search where the query is not a single sequence but a profile based on a multiple alignment of amino acid or nucleic acid sequences. The profile is used to query a sequence data set, i.e., a sequence database. The profile contains all the pertinent information for scoring each position in a sequence, in effect replacing the "scoring matrix" used for the standard query searches. The program used to query nucleotide databases with a protein profile is tprofilesearch. Tprofilesearch searches nucleic acid databases using an amino acid profile query. As the search is running, sequences in the database are translated to amino acid sequences in six reading frames. The output file for tprofilesearch is identical to the output file for profilesearch except for an additional column that indicates the frame in which the best alignment occurred.

The Smith-Waterman algorithm, (Smith and Waterman (1981) *supra*), is used to search for similarities between one sequence from the query and a group of sequences contained in the database. E score values as well as other sequence information, such as conserved peptide sequences of HXXXXD and PEG are used to identify related sequences. By using the conserved peptide sequence information, E score values of greater than E-12 and E-8 are considered. For example, the EST sequence originally used to identify ATAT2 had an E score of 0.0094, while the EST sequence originally used to identify ATLPAAT1 had an E score of 0.0868.

A protein sequence of glycerol-3-phosphate from *E. coli* (Swiss Prot Accession P00482) is used to search the NCBI non-redundant protein database using BLAST. In the first round of searches, other membrane forms of G3PAAT are identified. In subsequent PSI-BLAST searches (Altschul, *et al.* (1997) *Nucleic Acids Res* 25:3389-3402), LPAATs and other acyltransferases are identified. Using sequence alignment software programs, G3PAAT and different LPAAT amino acid sequences are aligned, and a profile is generated using a homologous sequence region, between amino acids 256 and 459 of the *E. coli* sequence.

The identified 204 amino acid is used to query the protein database using PSI-BLAST. After 5 iterations of PSI-BLAST, the profile generated from this new query (Figure 1)

identified soluble forms of G3PAAT. Prior to this identification, no sequence homology had been identified between the membrane and soluble forms of G3PAAT.

### 5    **Example 3: Excision of PSI-BLAST Profile**

The profile generated from the queries using PSI-BLAST is excised from the hyper text markup language (html) file. The worldwide web (www)/html interface to psiblast at ncbi stores the current generated profile matrix in a hidden field in the html file that is  
10    returned after each iteration of psiblast. However, this matrix has been encoded into string62 (s62) format for ease of transport through html. String62 format is a simple conversion of the values of the matrix into html legal ascii characters.

The encoded matrix width (x axis) is 26 characters, and comprise the consensus characters, the probabilities of each amino acid in the order A,B,C,D,E,F,G,H,I,K,L,M,N,  
15    P,Q,R,S,T,V,W,X,Y,Z (where B represents D and N, and Z represents Q and E, and X represents any amino acid), gap creation value, and gap extension value.

The length (y axis) of the matrix corresponds to the length of the sequences identified by PSI-BLAST. The order of the amino acids corresponds to the conserved amino acid sequence of the sequences identified using PSI-BLAST, with the N-terminal end at the top of  
20    the matrix. The probabilities of other amino acids at that position are represented for each amino acid along the x axis, below the respective single letter amino acid abbreviation.

Thus, each row of the profile consists of the highest scoring (consensus) amino acid, followed by the scores for each possible amino acid at that position in sequence matrix, the score for opening a gap at that position, and the score for continuing a gap at that position.

25    The string62 file is converted back into a profile for use in subsequent searches. The gap open field is set to 11 and the gap extension field is set to 1 along the x axis. The gap creation and gap extension values are known, based on the settings given to the PSI-BLAST algorithm. The matrix is exported to the standard GCG profile form. This format can be read by GenWeb.

30    The algorithm used to convert the string62 formatted file to the matrix is outlined in Table 1.

**Table 1**

1. if encoded character z then the value is blast score min
2. if encoded character Z then the value is blast score max
- 5 3. else if the encoded character is uppercase then its value is (64-(ascii # of char))
4. else if the encoded character is a digit the value is ((ascii # of char)-48)
5. else if the encoded character is not uppercase then the value is ((ascii # of char) - 87)
6. ALL B positions are set to min of D and N amino acids at that row in sequence matrix
7. ALL Z positions are set to min of Q and E amino acids at that row in sequence matrix
- 10 8. ALL X positions are set to min of all amino acids at that row in sequence matrix
9. kBLAST\_SCORE\_MAX=999;
10. kBLAST\_SCORE\_MIN=-999;
11. all gap opens are set to 11
12. all gap lens are set to 1

15

**Example 4: Identification of Novel Acyltransferase Related Amino Acid Sequences**

20 The profile (Figure 1) is used in further queries to identify a number of previously unidentified proteins from yeast as novel acyltransferases. A protein is identified from an *Arabidopsis* protein sequence database (ATAT1) (SEQ ID NO:2). Sequences are also identified from nucleic acid databases (Table 2)

25

**Table 2**

Database ID Number	BLAST Search Hits	Log probability
<u><i>Saccharomyces cerevisiae</i></u>		
gi 1078509	Limnanthes putative LPAAT	e-10 (SEQ ID
NO:217)		
30 gi 586485	Limnanthes putative LPAAT	e-13 (SEQ ID
NO:218)		

	gi 320748 NO:219)	Limnanthes putative LPAAT	e-19 (SEQ ID
	gi 2506920	SUPPRESSES CTR1 (choline transport mutant) (SEQ ID NO:220)	
5	gi 549627 NO:221)	similar to CTR1	e-118 (SEQ ID
	gi 2133031 NO:222)	unidentified	(SEQ ID
	gi 2132939 NO:223)	unidentified	(SEQ ID
10	gi 2132299 NO:224)	TFAZZIN	e-14 (SEQ ID

In Table 2, the gi number is the database identifier, the middle column shows the results of BLAST searches against the NCBI NR protein database, and the log probability number shows represents the log of the probability of such a match occurring by random chance. These proteins, including the ATAT1 protein sequence, are identified using the original PSI-BLAST search of the NCBI NR protein database. Thus, these proteins are novel acyltransferase related proteins with unidentified activities.

The *Arabidopsis* acyltransferase sequence, herein referred to as ATAT1, is also identified using the original PSI-BLAST search of the NCBI NR protein database, and did not have an annotated function.

Additional *Arabidopsis* amino acid sequences related to acyltransferases are identified from the databases, referred to as ATAT2est, ATAT3est, ATAT4est, ATAT5est, ATAT6est, ATAT7est, ATAT8est, ATAT9, ATAT10, and ATAT11est. Furthermore, *Arabidopsis* amino acid sequences are identified which demonstrate sequence similarity to known lysophosphatidic acid, referred to as ATLPAAT1. The sequences of ATAT9 and ATAT10 are identified from the database as genomic sequences, all other *Arabidopsis* sequences are identified as ESTs.

30

#### Example 5: Sequence Analysis of the Novel Acyltransferases

To obtain the entire coding region corresponding to the *Arabidopsis* acyltransferase sequences, synthetic oligo-nucleotide primers are designed to amplify the 5' and 3' ends of partial cDNA clones containing acyltransferase related sequences. Primers are designed according to the respective *Arabidopsis* acyltransferase related sequences (Table 3) and used  
5 in Rapid Amplification of cDNA Ends (RACE) reactions (Frohman *et al.* (1988) *Proc. Natl. Acad. Sci. USA* 85:8998-9002) using the Marathon cDNA amplification kit (Clontech Laboratories Inc, Palo Alto, CA). Primers with an R designation are used for 5' RACE reactions, and primers with an F designation are used for 3' RACE reactions.



Table 3

<u>ATAT2</u>		
ATAT2R1	CCATCCGCTTCAAGGGAACGACACCCATCA	(SEQ ID NO:135)
ATAT2R2	TCCCTGTCTTGCTTGATGAACTTAAAGCTTG	(SEQ ID NO:136)
5 ATAT2R3	ACAGCAGGAGTGTCTGATGATGGCAGATTC	(SEQ ID NO:137)
<u>ATAT3</u>		
ATAT3R1	ACTGGAGTTCCAGCCAAAAATGCACCTGTC	(SEQ ID NO:138)
ATAT3R2	GATACACCCTTGAAATCAGGCGATTTTGCT	(SEQ ID NO:139)
10	<u>ATAT4</u>	
ATAT4R1	TTGCAAATTCAATTCCTGTTTCACCGGGCC	(SEQ ID NO:140)
ATAT4R2	GTTTTCTGCTATTCCAGAAGGCGTCAACAA	(SEQ ID NO:141)
15	<u>ATAT5</u>	
ATAT5R1	CATTGAAGATCCGTCCGTGAAGTTNCCTTACC	(SEQ ID NO:142)
ATAT5R2	TCGAGCTGTGATCGATGATTGGCTGTGAAG	(SEQ ID NO:143)
ATAT5F1	GTCTCTTCAAAAACACACACACACGTCTCT	(SEQ ID NO:144)
ATAT5F2	GTCTCTTCAAAAACACACACACACGTCTCT	(SEQ ID NO:145)
20	<u>ATAT6</u>	
H76348-F1	GTAGAGAGCCTTACTTGCTTCGGTTTAGTC	(SEQ ID NO:146)
H76348-F2	ACGTCATCGTACCTGTTGCTATTGACTCAC	(SEQ ID NO:147)
H76348-R1	ACTTTTCCATTGTCAGGGACTCCTCGACAC	(SEQ ID NO:148)
25 H76348-R2	ACGGTGTAGGAAGGGAAAGGATTCAAAAGG	(SEQ ID NO:149)
<u>ATAT7</u>		
ATTS0193-F1	GCGATGAACTACAGAGTCGGATTCTTCCTC	(SEQ ID NO:150)
ATTS0193-F2	CCGGTTTACGAGATTACGTTCTTGAACCAG	(SEQ ID NO:151)
30 ATTS0193-R1	CAATGGAGACAAGGCTCGAAAGTGCTAACC	(SEQ ID NO:152)
ATTS0193-R2	ATTCTCTGAACATAGTTCGCCACGGTCATG	(SEQ ID NO:153)

ATAT8

AA042618-F1 GAAATCCAACGCCTTCCCAATATCACTCTG (SEQ ID NO:154)

AA042618-F2 CTTCAACTTTCCATCAGGATCTTGGCACGT (SEQ ID NO:155)

AA042618-R1 ACCACTTGTTAGAGACCTTACCTGCTTAGG (SEQ ID NO:156)

5 AA042618-R2 TCCTACCTACACCATCCAATTTCTCGACCC (SEQ ID NO:157)

ATAT11

ATAT11R1 CTGCGTCAAGTGAGCAACTCAGTTCTTGCA (SEQ ID NO:158)

ATAT11R2 TGGGAAGCAGCACGTTGTTTCAGTATCGGAA (SEQ ID NO:159)

10 ATAT11R3 TAGCCTCTGTGTAATCTGTGCCCTCGGGGA (SEQ ID NO:160)

From the nucleic acid sequences obtained from the RACE reactions, protein sequence is predicted for each nucleic acid sequence using Macvector software. Nucleic acid sequences are provided for ATAT1 (SEQ ID NO:1), ATAT2 (SEQ ID NO:3), ATAT3 (SEQ ID NO:5), ATAT4 (SEQ ID NO:7), ATAT5 (SEQ ID NO:9), ATAT6 (SEQ ID NO:10), ATAT7 (SEQ ID NO:12), ATAT8 (SEQ ID NO:14), ATAT9 (SEQ ID NO:16), ATAT10 (SEQ ID NO:18), ATAT11 (SEQ ID NO:20) and ATLPAAT1 (SEQ ID NO:22), respectively.

The protein sequence derived from the ATAT1 (SEQ ID NO:2) nucleic acid sequence from Arabidopsis has a predicted molecular mass of 32.5 kDa, and a PI of 9.74. Alignment of the Arabidopsis acyltransferase with several LPAAT and G3PAAT shows that some of the domains that are conserved between LPAAT and G3PAAT are conserved in the new acyltransferase protein.

The ATAT2 nucleic acid sequence is predicted to encode a 312 amino acid protein (SEQ ID NO:4), with a molecular weight of 34.6 kD, and a pI of 9.99. The ATAT2 protein may also contain 2 to 3 transmembrane domains. However, the protein encoded by the ATAT2 nucleic acid sequence may be longer than predicted because of the absence of an inframe stop codon upstream of the ATG start codon used.

The ATAT3 nucleic acid sequence is predicted to encode a 398 amino acid protein (SEQ ID NO:6), with a molecular weight of 44.7 kD, and a pI of 5.62. The ATAT3 protein may contain 1 to 4 transmembrane domains. The ATAT4 nucleic acid sequence is predicted to encode a 317 amino acid protein (SEQ ID NO:8), with a molecular weight of 36.5 kD, and a pI of 9.67. The ATAT4 protein is predicted to have 2 to 5 transmembrane domains.

The ATLPAAT1 nucleic acid sequence is predicted to encode a 389 amino acid protein (SEQ ID NO:23), with a molecular weight of 43.7 kD, and a pI of 9.52. The ATLPAAT1 protein is predicted to have up to 3 transmembrane domains. The protein predicted from the ATLPAAT1 nucleic acid sequence is similar to LPAATs reported for *Brassica*, maize, and meadowfoam (described in PCT Publication WO 94/13814). The ATAT11 nucleic acid sequence is predicted to encode a 375 amino acid protein (SEQ ID NO:21), with a molecular weight of 43.5 kD, and a pI of 9.45. The deduced amino acid sequences of ATAT6 (SEQ ID NO:11), ATAT7 (SEQ ID NO:13), ATAT8 (SEQ ID NO:15), ATAT9 (SEQ ID NO:17), and ATAT10 (SEQ ID NO:19) are also provided

A sequence region approximately 30 amino acids upstream through approximately 100 amino acids downstream of the conserved amino acid sequences HXXXXD (Heath and Rock, (1998) *J. Bacteriol.* 180(6):1425-1430) and PEG (Neuwald (1997) *Curr Biol* 7:R465-R466) of the predicted amino acid sequences derived from the nucleic acid sequences of ATAT1, ATAT2, ATAT3, ATAT4, ATAT6, ATAT7, ATAT8, ATAT9, ATAT10, ATLPAAT1, and ATAT11 are compared to the amino acid sequences of lysophosphatidic acid acyltransferase (Jojoba AT (SEQ ID NO:162, the nucleic acid sequence is provided in SEQ ID NO:161), maize AT (PCT Publication WO 94/13814), PLSC coco(GenBank accession 1098605), PLSC Lim(GenBank accession 1209507), PLSC, Ecoli (GenBank accession 1209507), and PLSC Yeast(GenBank accession 464422)) and glycerol-3-phosphate acyltransferase (PLSB Ecoli(GenBank accession 130326) and PLSB Mouse(GenBank accession 2498786)) (Figure 2), and similarities are identified (Figure 2 and Figure 3).

Sequence comparisons reveal several classes of acyltransferases exist based on conserved amino acid sequences identified in the comparisons in Figure 2. For example, ATAT1, ATAT6, ATAT7, ATAT8, and ATAT9, contain the conserved amino acid sequences of VTYSXS(SEQ ID NO: 128), VXLTRXR(SEQ ID NO: 129), LXXGDLV(SEQ ID NO: 132) between the HXXXXD and PEG sequences. In addition, ATAT1, ATAT6, ATAT7, ATAT8, and ATAT9 also contain the conserved sequences CPEGT(SEQ ID NO: 130) which comprises the PEG sequence, as well as IVPVA(SEQ ID NO: 131) and VANXXQ (SEQ ID NO: 134)(Figure 2) downstream of the PEG sequence. The sequences corresponding to ATAT1, ATAT7, and ATAT9 are the most closely related in this class, with similarities between ATAT1 and ATAT9 of 67.0%, between ATAT1 and ATAT7 of 58.2% and between ATAT9 and ATAT7 of 63.9% (Figure 3B).

Sequence comparisons also demonstrate that the sequence of ATLPAAT1 is most closely related to the jojoba LPAAT (82.3% similar), and maize (78.0% similar).

Furthermore, sequence analysis demonstrates that ATAT4 is the most divergent sequence with the highest similarity to ATAT10 (18.5%). The highest similarity (15.3%) to a known sequence is with a meadowfoam (*Limnanthes douglassi*) LPAAT. However, the sequences of ATAT4 and ATAT10 share several conserved peptide sequences with the amino acid sequences of ATAT2 and ATAT3 (Figure 2), VXNHXS (SEQ ID NO: 127) where the H comprises the conserved H of the HXXXXD sequence and FXXGAF (SEQ ID NO: 133) downstream of the PEG sequence.

#### Example 6: Identification of Additional Acyltransferase Sequences

The novel *Arabidopsis* sequences identified above are used to search proprietary databases containing soybean and corn EST sequences. The results of this search identifies EST sequences from soybean (SEQ ID NO:24 through SEQ ID NO: 85) as well as from corn (SEQ ID NO: 86 through SEQ ID NO:126) as encoding acyltransferase related proteins.

Sequence comparisons between the various EST sequences and the complete *Arabidopsis* sequences reveals that the identified EST sequences demonstrate higher similarity to the various *Arabidopsis* sequences as determined by BLAST scores.

Expressed Sequence Tag (EST) sequences from soybean and corn databases are identified which are most closely related by BLAST score to ATAT1 (SEQ ID NOS:24-29 and SEQ ID NOS:86-88, respectively), ATAT2 (SEQ ID NO: 30 and SEQ ID NO:89, respectively), ATAT3 (SEQ ID NOS:31-35 and SEQ ID NOS:90-94, respectively), ATAT4 (SEQ ID NOS:36-44 and SEQ ID NOS:95-100, respectively), ATAT6 (SEQ ID NOS:45-49 and SEQ ID NO:101, respectively), ATAT7 (SEQ ID NOS:50-54 and SEQ ID NOS:102-103, respectively), ATAT8 (SEQ ID NOS:55-56 and SEQ ID NO:104, respectively), ATAT9 (SEQ ID NOS:57-79 and SEQ ID NOS:105-111, respectively), ATAT10 (SEQ ID NOS:80-81 and SEQ ID NO:112, respectively), ATAT11, (SEQ ID NOS:82-85 and SEQ ID NOS:123-126, respectively), and ATLPAAT1 (SEQ ID NOS: 113-122 respectively).

**Example 7: Expression Construct Preparation**

A series of synthetic oligo nucleotide primers were prepared for use in Polymerase Chain Reactions (PCR) to amplify the entire DNA sequences encoding the various acyltransferase sequences identified above. The sequences are listed in Table 3.

**Table 3**

<b>Primer</b>	<b>Sequence (listed 5'-3')</b>	<b>SEQ ID NO:</b>
ATAT1F	AAGCTTGCATGCGTCGACACAATGGTTCATGCGACCAAGT CAG	163
ATAT1R	GGTACCGTCGACTCACTTCTTGGTGTGTTGATAG	164
ATAT2F	GGATCCGCGGCCGCACAATGACGAGCTTTACTACTTCCCT TCAT	165
ATAT2R	GGATCCCCCTGCAGGTTAGAGATCCATTGATTCTGCAAT	166
ATAT3F	GGATCCGCGGCCGCATAATGGAATCAGAGCTCAAAGAT	167
ATAT3R	GGATCCCCCTGCAGGTCATTCTTCTTTCTGATGGAAATC	168
ATAT4F	GGATCCGCGGCCGCACAATGACTCGTTCACAAGATGTTTC A	169
ATAT4R	GGATCCCCCTGCAGGTCAGTCTCTTCCAATCTAGCCAG	170
ATAT6F	GGATCCGCGGCCGCACAATGTCCGGTAATAAGATCTCGAC TCTTCA	171
ATAT6R	GGATCCCCCTGCAGGTTATTTTTTCTTGACAACTCCGTTAT TACCGG	172
ATAT7F	ATATCCGCGGCCGCACAATGGTTATGGAGCAAGCTGGAA	173
ATAT7R	GGATCCCCCTGCAGGTCAATGGAGACAAGGCTCGAAAGT	174
ATAT8F	GGATCCGCGGCCGCACAATGTCCGCCAAGATTTCAATATT CC	175
ATAT8R	GGATCCCCCTGCAGGTTAATTTTTCTTAATACTACTCCATT	176
ATAT9F	GGATCCGCGGCCGCACAATGGGAGCTCAGGAGAAACGGCG CC	177
ATAT9R	GGATCCCCCTGCAGGTCACGTCTTCTCCTTCTTCACCGG	178
ATAT10F	GGATCCGCGGCCGCACAATGGCGGATCCTGATCTGTCTTC TCCT	179
ATAT10R	GGATCCCCCTGCAGGTTATGTTGGGGCCAAGTCAGGTGCAA AGAT	180
ATAT11F	GGATCCGCGGCCGCAAAATGGAAAAAAGAGTGTAACAAA	181

	TTCT	
ATAT11R	GGATCCCCTGCAGGTTATTTGTTTACTAATTTGAGGGAAT	182
	TTTTTG	
ATLPAAT	TCGACCTGCAGGAAGCTTAAGGATGGTGATTGCTGC	183
1F		
ATLPAAT	GGATCCGCGGCCGCTTACTTCTCCTTCTCCG	184
1R		
YSCAT1F	GGATCCGCGGCCGCGACAATGTCTTTTAGGGATGTCCTAG	185
YSCAT1R	GGATCCCCTGCAGGTCAATCATCCTTACCCTTTGGTTTAC	186
	C	
YSCAT 1	ATGTCTTTTAGGGATGTCCTAGAAAGAGGAGATGAATTTT	187
KO F	CTGTGCGGTATTTTCACACCG	
YSCAT 1	TCAATCATCCTTACCCTTTGGTTTACCCTCTGGAGGCAGA	188
KO R	AGATTGTACTGAGAGTGCAC	
YSCAT2F	GGATCCGCGGCCGCGACAATGAAGCATTCCCAAAAATACCG	189
	TAGG	
YSCAT2R	GGATCCCCTGCAGGTCAATGATTTTTTTTCATCACAAATA	190
	C	
YSCAT 2	ATGAAGCATTCCCAAAAATACCGTAGGTATGGAATTTATG	191
KO F	CTGTGCGGTATTTTCACACCG	
YSCAT 2	TCAATGATTTTTTTTCATCACAAATACAAGAATAAGAAAA	192
KO R	AGATTGTACTGAGAGTGCAC	
YSCAT	GGATCCGCGGCCGCGACAATGGGTTTTGTTGATTTCTTCGA	193
3F	AAC	
YSCAT	GGATCCCCTGCAGGTTATTTGGTCTCAATTTTAATATTTT	194
3R	TTTGC	
YSCAT 3	ATGGGTTTTGTTGATTTCTTCGAAACATATATGGTCGGTT	195
KO F	CTGTGCGGTATTTTCACACCG	
YSCAT 3	TTATTTGGTCTCAATTTTAATATTTTTTTTGCAAGGACTCG	196
KO R	AGATTGTACTGAGAGTGCAC	
YSCAT	GGATCCGCGGCCGCGACAATGGAAAAGTACACCAATTGGAG	197
4F	AGAC	
YSCAT	GGATCCCCTGCAGGCTACTTCCTCTTTTACGTTGATCGC	198
4R	TG	
YSCAT 4	ATGGAAAAGTACACCAATTGGAGAGACAATGGTACGGGAA	199
KO F	CTGTGCGGTATTTTCACACCG	
YSCAT 4	CTACTTCCTCTTTTTACGTTGATCGCTGATATATTCCTTC	200
KO R	AGATTGTACTGAGAGTGCAC	

YSCAT	GGATCCGCGGCCGCACAATGCCTGCACCAAACTCACGGA	201
5F	G	
YSCAT	GGATCCCCTGCAGGCTACGCATCTCCTTCTTTCCCTTC	202
5R		
YSCAT 5	ATGCCTGCACCAAACTCACGGAGAAATCTGCCTCTTCCA	203
KO F	CTGTGCGGTATTTACACCG	
YSCAT 5	CTACGCATCTCCTTCTTTCCCTTCTTCTTCTTCTCCTCT	204
KO R	AGATTGTACTGAGAGTGCAC	
YSCAT	GGATCCGCGGCCGCACAATGTCTGCTCCCGCTGCCGATCA	205
6F	TAACGC	
YSCAT	GGATCCCCTGCAGGTCATTCTTTCTTTTCGTGTTCTCTTT	206
6R	TCTG	
YSCAT 6	ATGTCTGCTCCCGCTGCCGATCATAACGCTGCCAAACCTA	207
KO F	CTGTGCGGTATTTACACCG	
YSCAT 6	TCATTCTTTCTTTTCGTGTTCTCTTTTCTGTCTTACCAGC	208
KO R	AGATTGTACTGAGAGTGCAC	
YSCAT	GGATCCGCGGCCGCACAATGCTGCATCAAAAAATAGCTCA	209
7F	TAAAGTTCG	
YSCAT	GGATCCCCTGCAGGTCAAAAAATAAAACAATAAAGTTTAT	210
7R	AAACTAACC	
YSCAT 7	ATGCTGCATCAAAAAATAGCTCATAAAGTTCGAAAAGTCG	211
KO F	CTGTGCGGTATTTACACCG	
YSCAT 7	TCAAAAAATAAAACAATAAAGTTTATAAACTAACCAAATT	212
KO R	AGATTGTACTGAGAGTGCAC	
YSCAT	GGATCCGCGGCCGCACAATGAGTGTGATAGGTAGGTTCTT	213
8F	G	
YSCAT	GGATCCCCTGCAGGTTAATGCATCTTTTTTACAGATGAAC	214
8R	C	
YSCAT 8	ATGAGTGTGATAGGTAGGTTCTTGTATTACTTGAGGTCCG	215
KO F	CTGTGCGGTATTTACACCG	
YSCAT 8	TTAATGCATCTTTTTTACAGATGAACCTTCGTTATGGGTA	216
KO R	AGATTGTACTGAGAGTGCAC	

The entire coding regions for each of the acyltransferase sequences were amplified using the respective primers listed in the Table 3 above, cloned into the vector pCR2.1Topo (Invitrogen) or pZero (Invitrogen), and labeled as pCGN8558 (ATAT1), pCGN8564

(ATAT2), pCGB8565 (ATAT3), pCGN8566 (ATAT4), pCGN8918 (ATAT6), pCGN8913 (ATAT7), pCGN8904 (ATAT8), pCGN9970 (ATAT9), pCGN9940 (ATAT10), pCGN8567 (ATAT11), pCGN8632 (ATLPAAT1), pCGN9901 (YSCAT1 also referred to as gi2132299), pCGN9902 (YSCAT2, also referred to as gi1078509), pCGN9903 (YSCAT3, also referred to as gi2132939), pCGN9904 (YSCAT4, also referred to gi2133031), pCGN9905 (YSCAT5, also referred to as gi320748), pCGN9906 (YSCAT6, also referred to as gi549627), pCGN9907 (YSCAT7, also referred to as gi586485), and pCGN9908 (YSCAT8, also referred to as gi464422). The nucleic acid sequences for the respective yeast acyltransferase are provided YSCAT1 (SEQ ID NO:225), YSCAT2 (SEQ ID NO:226), YSCAT3 (SEQ ID NO:227), YSCAT4 (SEQ ID NO:228), YSCAT5 (SEQ ID NO:229), YSCAT6 (SEQ ID NO:230), YSCAT7 (SEQ ID NO:231), and YSCAT8 (SEQ ID NO:232).



#### 7A. Baculovirus Expression Constructs

Constructs are prepared to direct the expression of the *Arabidopsis* ATAT sequences in cultured insect cells. The entire coding regions of ATAT1, 2, 3, 4, 6, 7, 8, 9, 10, and 11 are cloned into the vector pFastBac1 (Gibco-BRL, Gaithersburg, MD) digested with *NotI* and *PstI*. The respective coding sequences were cloned as *NotI/Sse8387I* fragments. Double stranded DNA sequence was obtained to verify that no errors were introduced by PCR amplification. The resulting plasmid were designated pCGN9723 (ATAT1), pCGN9724 (ATAT2), pCGN9725 (ATAT3), pCGN9726 (ATAT4), pCGN9727 (ATAT5), pCGN9728 (ATAT7), pCGN9729 (ATAT8), pCGN9991 (ATAT9) pCGN9730 (ATAT10), pCGN9731 (ATAT11).

#### 7B. Plant Expression Construct Preparation

A plasmid containing the napin cassette derived from pCGN3223 (described in USPN 5,639,790, the entirety of which is incorporated herein by reference) was modified to make it more useful for cloning large DNA fragments containing multiple restriction sites, and to allow the cloning of multiple napin fusion genes into plant binary transformation vectors. An adapter comprised of the self annealed oligonucleotide of sequence CGCGATTTAAATGGCGCGCCCTGCAGGCGGCCGCTGCAGGGCGCGCCATTTAA (SEQ ID NO:233) AT was ligated into the cloning vector pBC SK+ (Stratagene) after digestion with the restriction endonuclease BssHII to construct vector pCGN7765. Plasmids pCGN3223 and pCGN7765 were digested with *NotI* and ligated together. The resultant vector, pCGN7770, contains the pCGN7765 backbone with the napin seed specific expression cassette from pCGN3223.

The cloning cassette, pCGN7787, essentially the same regulatory elements as pCGN7770, with the exception of the napin regulatory regions of pCGN7770 have been replaced with the double CAMV 35S promoter and the tml polyadenylation and transcriptional termination region.

A binary vector for plant transformation, pCGN5139, was constructed from pCGN1558 (McBride and Summerfelt, (1990) Plant Molecular Biology, 14:269-276). The polylinker of pCGN1558 was replaced as a *HindIII*/*Asp718* fragment with a polylinker containing unique restriction endonuclease sites, *AscI*, *PacI*, *XbaI*, *SwaI*, *BamHI*, and *NotI*. The *Asp718* and *HindIII* restriction endonuclease sites are retained in pCGN5139.

A series of turbo binary vectors are constructed to allow for the rapid cloning of DNA sequences into binary vectors containing transcriptional initiation regions (promoters) and transcriptional termination regions.

The plasmid pCGN8618 was constructed by ligating oligonucleotides 5'-  
5 TCGAGGATCCGCGGCCGCAAGCTTCCTGCAGG-3' ) (SEQ ID NO:234) and 5'-  
TCGACCTGCAGGAAGCTTGCGGCCGCGGATCC-3' ) (SEQ ID NO:235) into SalI/XhoI-  
digested pCGN7770. A fragment containing the napin promoter, polylinker and napin 3'  
region was excised from pCGN8618 by digestion with Asp718I; the fragment was blunt-  
ended by filling in the 5' overhangs with Klenow fragment then ligated into pCGN5139 that  
10 had been digested with Asp718I and HindIII and blunt-ended by filling in the 5' overhangs  
with Klenow fragment. A plasmid containing the insert oriented so that the napin promoter  
was closest to the blunted Asp718I site of pCGN5139 and the napin 3' was closest to the  
blunted HindIII site was subjected to sequence analysis to confirm both the insert orientation  
and the integrity of cloning junctions. The resulting plasmid was designated pCGN8622.

15 The plasmid pCGN8619 was constructed by ligating oligonucleotides 5'-  
TCGACCTGCAGGAAGCTTGCGGCCGCGGATCC -3' ) (SEQ ID NO:236) and 5'-  
TCGAGGATCCGCGGCCGCAAGCTTCCTGCAGG-3' ) (SEQ ID NO:237) into SalI/XhoI-  
digested pCGN7770. A fragment containing the napin promoter, polylinker and napin 3'  
region was removed from pCGN8619 by digestion with Asp718I; the fragment was blunt-  
20 ended by filling in the 5' overhangs with Klenow fragment then ligated into pCGN5139 that  
had been digested with Asp718I and HindIII and blunt-ended by filling in the 5' overhangs  
with Klenow fragment. A plasmid containing the insert oriented so that the napin promoter  
was closest to the blunted Asp718I site of pCGN5139 and the napin 3' was closest to the  
blunted HindIII site was subjected to sequence analysis to confirm both the insert orientation  
25 and the integrity of cloning junctions. The resulting plasmid was designated pCGN8623.

The plasmid pCGN8620 was constructed by ligating oligonucleotides 5'-  
TCGAGGATCCGCGGCCGCAAGCTTCCTGCAGGAGCT -3' ) (SEQ ID NO:238) and 5'-  
CCTGCAGGAAGCTTGCGGCCGCGGATCC-3' ) (SEQ ID NO:239) into SalI/SacI-  
digested pCGN7787. A fragment containing the d35S promoter, polylinker and tml 3' region  
30 was removed from pCGN8620 by complete digestion with Asp718I and partial digestion with  
NotI. The fragment was blunt-ended by filling in the 5' overhangs with Klenow fragment  
then ligated into pCGN5139 that had been digested with Asp718I and HindIII and blunt-  
ended by filling in the 5' overhangs with Klenow fragment. A plasmid containing the insert

oriented so that the d35S promoter was closest to the blunted Asp718I site of pCGN5139 and the tml 3' was closest to the blunted HindIII site was subjected to sequence analysis to confirm both the insert orientation and the integrity of cloning junctions. The resulting plasmid was designated pCGN8624.

- 5           The plasmid pCGN8621 was constructed by ligating oligonucleotides 5'-TCGACCTGCAGGAAGCTTGCGGCCGCGGATCCAGCT -3' ) (SEQ ID NO:240) and 5'-GGATCCGCGGCCGCAAGCTTCCTGCAGG-3' ) (SEQ ID NO:241) into SalI/SacI-digested pCGN7787. A fragment containing the d35S promoter, polylinker and tml 3' region was removed from pCGN8621 by complete digestion with Asp718I and partial digestion with
- 10   NotI. The fragment was blunt-ended by filling in the 5' overhangs with Klenow fragment then ligated into pCGN5139 that had been digested with Asp718I and HindIII and blunt-ended by filling in the 5' overhangs with Klenow fragment. A plasmid containing the insert oriented so that the d35S promoter was closest to the blunted Asp718I site of pCGN5139 and the tml 3' was closest to the blunted HindIII site was subjected to sequence analysis to
- 15   confirm both the insert orientation and the integrity of cloning junctions. The resulting plasmid was designated pCGN8625.

- The coding regions of the various acyltransferase sequences were cloned as *NotI/Sse8387I* fragments into pCGN8622, pCGN8623, pCGN8624, and pCGN8625, for expression in sense or antisense orientations from a tissue preferential promoter, napin, or the
- 20   35S promoter. Fragments which were cloned into the pCGN8622 vector created the constructs pCGN8901 (ATAT1), pCGN8571 (ATAT2), pCGN8909 (ATAT3), pCGN8596 (ATAT4), pCGN8919 (ATAT6), pCGN8914 (ATAT7), pCGN8905 (ATAT8), pCGN9973 (ATAT9), pCGN9942 (ATAT10), pCGN8575 (ATAT11), and pCGN8633 (ATLPAAT1) for the sense expression of the respective coding sequences from the napin promoter. Fragments
- 25   which were cloned into the pCGN8623 vector created the constructs pCGN8900 (ATAT1), pCGN8572 (ATAT2), pCGN8910 (ATAT3), pCGN8597 (ATAT4), pCGN8920 (ATAT6), pCGN8915 (ATAT7), pCGN8906 (ATAT8), pCGN9972 (ATAT9), pCGN9943 (ATAT10), pCGN8576 (ATAT11), and pCGN8634 (ATLPAAT1) for the antisense expression of the respective coding sequences from the napin promoter. Fragments which were cloned into the
- 30   pCGN8624 vector created the constructs pCGN8903 (ATAT1), pCGN8573 (ATAT2), pCGN8911 (ATAT3), pCGN8598 (ATAT4), pCGN8921 (ATAT6), pCGN8916 (ATAT7), pCGN8907 (ATAT8), pCGN9971 (ATAT9), pCGN9944 (ATAT10), pCGN8577 (ATAT11), and pCGN8635 (ATLPAAT1) for the sense expression of the respective coding sequences

from the 35S promoter. Fragments which were cloned into the pCGN8625 vector created the constructs pCGN8902 (ATAT1) and pCGN9974 (ATAT9) for the antisense expression of the respective coding sequences from the 35S promoter.

In addition, the yeast acyltransferase coding sequences were cloned into the vector pCGN8624 creating the constructs pCGN9926 (YSCAT1), pCGN9927 (YSCAT2), pCGN9928 (YSCAT3), pCGN9929 (YSCAT4), pCGN9930 (YSCAT5), pCGN9931 (YSCAT6), pCGN9932 (YSCAT7), and pCGN9933 (YSCAT8). These constructs allow for the sense expression of the respective acyltransferase coding sequences from the 35S promoter in plant cells.

#### Example 8: Plant Transformation

A variety of methods have been developed to insert a DNA sequence of interest into the genome of a plant host to obtain the transcription or transcription and translation of the sequence to effect phenotypic changes.

Transgenic *Brassica* plants are obtained by *Agrobacterium*-mediated transformation as described by Radke *et al.* (*Theor. Appl. Genet.* (1988) 75:685-694; *Plant Cell Reports* (1992) 11:499-505). Transgenic *Arabidopsis thaliana* plants may be obtained by *Agrobacterium*-mediated transformation as described by Valverkens *et al.*, (*Proc. Nat. Acad. Sci.* (1988) 85:5536-5540), or as described by Bent *et al.* ((1994), *Science* 265:1856-1860), or Bechtold *et al.* ((1993), *C.R.Acad.Sci, Life Sciences* 316:1194-1199) or Clough, *et al.* (1998) *Plant J.*, 16:735-43. Other plant species may be similarly transformed using related techniques.

Alternatively, microprojectile bombardment methods, such as described by Klein *et al.* (*Bio/Technology* 10:286-291) may also be used to obtain nuclear transformed plants.

The above results demonstrate that the nucleic acid sequences identified encode proteins which are related to protein sequences encoding acyltransferase proteins. Such acyltransferase sequences find use in preparing expression constructs for plant transformations.

All publications and patent applications mentioned in this specification are indicative of the level of skill of those skilled in the art to which this invention pertains. All

publications and patent applications are herein incorporated by reference to the same extent as if each individual publication or patent application was specifically and individually indicated to be incorporated by reference.

Although the foregoing invention has been described in some detail by way of  
5 illustration and example for purposes of clarity of understanding, it will be obvious that certain changes and modifications may be practiced within the scope of the appended claim.

### Claims

What is Claimed is:

1. An isolated DNA sequence encoding an enzyme of the class of acyltransferase-like  
5 proteins,

wherein said enzyme includes the amino acid sequence of SEQ ID NO: 127  
(VxNHxS) wherein the H is the conserved Histidine residue in the conserved peptide  
sequence HXXXXD of said acyltransferase-like protein, x representing any amino acid.

- 10 2. An isolated DNA sequence encoding an enzyme of the class of acyltransferase-like  
proteins,

wherein said enzyme includes the amino acid sequence of SEQ ID NO: 128  
(VTYSxS) within about 30 amino acids downstream from the conserved amino acid sequence  
HXXXXD of said acyltransferase-like protein, x representing any amino acid.

- 15 3. An isolated DNA sequence encoding an enzyme of the class of acyltransferase-like  
proteins,

wherein said enzyme includes the amino acid sequence of SEQ ID NO: 129  
(VxLTRxR) within about 60 amino acids downstream from the conserved amino acid  
20 sequence HXXXXD of said acyltransferase-like protein, x representing any amino acid.

4. An isolated DNA sequence encoding an enzyme of the class of acyltransferase-like  
proteins,

wherein said enzyme includes the amino acid sequence of SEQ ID NO: 132  
25 (LxxGDLV) within about 20 amino acids upstream of the conserved amino acid sequence  
PEG of said acyltransferase-like protein, x representing any amino acid.

5. An isolated DNA sequence encoding an enzyme of the class of acyltransferase-like  
proteins,

30 wherein said enzyme includes the amino acid sequence of SEQ ID NO: 130 (CPEGT)  
containing the conserved amino acid sequence PEG of said acyltransferase-like protein.

6. An isolated DNA sequence encoding an enzyme of the class of acyltransferase-like proteins,

wherein said enzyme includes the amino acid sequence of SEQ ID NO: 133 (FxxGAF) within about 20 amino acids downstream from the conserved amino acid sequence

5 PEG of said acyltransferase-like protein, x representing any amino acid.

7. An isolated DNA sequence encoding an enzyme of the class of acyltransferase-like proteins,

wherein said enzyme includes the amino acid sequence of SEQ ID NO: 131 (IVPVA) within about 40 amino acids downstream from the conserved amino acid sequence PEG of said acyltransferase-like protein.

8. An isolated DNA sequence encoding an enzyme of the class of acyltransferase-like proteins,

15 wherein said enzyme includes the amino acid sequence of SEQ ID NO: 134 (VANxxQ) within about 110 amino acids downstream from the conserved amino acid sequence PEG of said acyltransferase-like protein, x representing any amino acid.

20 9. A DNA sequence encoding an enzyme of the class of acyltransferase-like proteins, said DNA sequence obtainable by the steps comprising:

(a) using the profile of Figure 1 to search a nucleic acid sequence database;

(b) obtaining a probability score for nucleic acid sequences in said sequence database using the Smith-Waterman algorithm; and

25 (c) selecting a nucleic acid sequence having a probability score of less than about 1.

10. The DNA encoding sequence according to Claim 9, wherein said DNA sequence is an encoding sequence.

30 11. The DNA encoding sequence according to Claim 9, wherein said DNA sequence is an EST.

12. The DNA encoding sequence according to any one of Claims 1 to 11, wherein said acyltransferase-like protein is from a plant.

13. A construct comprising a DNA sequence of any one of Claims 1 to 11 linked to a heterologous transcriptional and translational initiation region functional in a host cell.

14. The construct according to Claim 13 wherein said host cell is a plant cell.

15. A plant cell comprising a DNA construct according to Claim 13.

16. A plant comprising a cell according to Claim 15.

17. The DNA encoding sequence of any one of 1 to 11 wherein said acyltransferase-like protein is from *Arabidopsis thaliana*.

18. The DNA encoding sequence of any one of 1 to 11 wherein said acyltransferase-like protein is from corn.

19. The DNA encoding sequence of Claim 18 wherein said sequence comprises and EST selected from the group consisting of SEQ ID NO: 86 through SEQ ID NO: 126.

20. The DNA encoding sequence of any one of 1 to 11 wherein said acyltransferase-like protein is from soybean.

21. The DNA encoding sequence of Claim 20 wherein said sequence comprises and EST selected from the group consisting of SEQ ID NO: 24 through SEQ ID NO: 85.

22. The DNA encoding sequence of any one of Claims 2, 3, 4, 5, 7 and 8 wherein said acyltransferase-like protein is selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14 and SEQ ID NO: 16.



23 . The DNA encoding sequence of either of Claim 1 and Claim 6 wherein said acyltransferase-like protein is selected from the group consisting of SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7 and SEQ ID NO: 18.

Con	A	B	C	D	E	F	G	H	I	K	L	M	N	P	Q	R	S	T	V	W	X	Y	Z	Gap	Len
S	0	0	-1	0	0	-2	0	-1	-2	0	-2	-1	0	-1	0	-1	5	0	-2	-3	-3	-3	0	11	1
K	0	-1	-3	-1	0	-3	-2	-1	-3	5	-2	-1	0	-2	0	0	2	0	-2	-4	-4	-2	0	11	1
Q	-2	-1	-3	-1	0	-1	-3	-2	-2	3	1	0	2	-2	4	0	-1	-1	-2	-4	-4	-2	0	11	1
I	0	-1	-2	-1	1	-3	3	-2	5	0	0	0	0	-3	-1	-2	-2	-1	1	-4	-4	-2	-1	11	1
G	0	-1	-3	-1	1	-3	3	-2	5	-1	0	-2	0	-2	-1	-2	3	2	-3	-3	-3	-3	-1	11	1
I	0	-1	-3	-1	1	-3	3	-2	5	-1	0	-2	0	-2	-2	-2	0	-1	0	-4	-4	-2	-2	11	1
N	-2	-1	-4	-1	2	-4	-2	-1	-4	4	-3	-2	4	1	2	0	1	-1	-4	-4	-4	-3	1	11	1
K	-2	-1	-4	-1	2	-4	-2	-1	-4	4	-3	-2	4	1	2	0	1	-1	-4	-4	-4	-3	-2	11	1
T	1	-3	-2	-3	-2	-1	-3	-2	0	-2	-1	-1	3	-2	1	-2	1	3	2	-2	-3	3	-2	11	1
10	0	-1	-4	-1	2	-4	-2	-1	-4	5	-3	-2	-1	-2	4	2	-1	-2	-3	-4	-4	2	2	11	1
K	2	-2	-3	-2	-1	-4	-2	-2	-4	3	-3	-2	1	-2	1	3	1	-2	-3	-4	-4	-3	-1	11	1
K	0	-1	-4	-1	4	-3	-3	-2	0	2	1	-2	0	-3	0	1	1	-1	-3	-4	-4	-3	-1	11	1
E	3	-4	-2	-4	-3	5	-3	-3	0	3	-2	-2	0	-3	0	-3	-2	-2	0	-2	-4	0	-3	11	1
F	0	-2	-4	-2	1	4	-5	-2	2	-4	3	-2	3	-5	-3	-4	0	-2	2	4	-4	-2	-4	11	1
F	-3	-5	-4	-5	-4	0	-4	-4	0	-3	3	1	-4	-5	-3	-4	0	-2	0	-5	-5	-4	-4	11	1
W	-2	-4	-3	-5	-4	0	-4	-4	0	-3	3	1	-4	-5	-3	-4	0	-2	2	-4	-4	-2	-4	11	1
P	-2	-4	-3	-5	-4	0	-4	-4	0	-3	3	1	-4	-5	-3	-4	0	-2	2	-4	-4	-2	-4	11	1
E	-2	-2	-4	-1	5	-4	-3	-2	-4	1	-3	2	-2	0	1	-2	0	0	-3	-5	-5	-4	1	11	1
20	-2	-5	-2	-5	-4	-2	-5	-2	6	-4	1	5	-4	-4	-3	-4	-3	-2	1	-4	-5	-2	-4	11	1
I	3	-3	-2	-3	0	-3	-3	-3	-1	-2	1	3	-3	-3	-2	-4	1	-2	0	-4	-4	-3	-2	11	1
A	4	0	-3	1	-1	1	-2	2	-3	0	-3	2	0	0	2	-2	-1	-2	-3	-4	-4	-3	-2	11	1
A	-2	1	-4	1	-1	1	-4	1	0	2	-3	1	0	-3	1	4	0	0	-3	-5	-5	-3	-1	11	1
L	0	-5	-3	-5	-4	0	-5	-4	0	0	4	0	-4	-4	-3	0	2	0	4	-4	-4	0	-4	11	1
S	1	-2	-3	-5	-4	0	-3	0	-3	1	-4	-3	2	5	1	1	-1	0	-4	-3	-5	4	0	11	1
P	-3	-2	-4	-4	-1	5	0	3	2	-2	0	3	-3	-5	-2	3	0	-3	-3	6	-5	-1	-2	11	1
W	0	-4	-3	-4	-3	-2	-4	-3	2	-3	3	4	2	-4	0	-4	0	-2	-1	4	-4	-3	-4	11	1
C	-3	-4	-3	-4	-4	1	-5	-3	2	-4	1	-1	1	-5	-4	-4	-1	0	3	-3	-5	3	-4	11	1
30	0	-3	-4	-3	-2	-4	-3	-2	0	2	-3	-3	0	0	0	0	1	0	-3	-5	-5	0	-2	11	1
R	0	-5	1	-5	-4	0	-5	-4	1	-3	0	5	-4	-4	-3	4	0	0	1	7	-5	-3	-4	11	1
M	-1	-5	6	-5	-4	2	-1	-4	1	-2	0	0	-4	0	-4	3	-3	3	0	4	-5	0	-4	11	1
W	-3	-2	-4	-4	0	-2	0	-3	-3	0	-1	-3	-1	0	-2	-2	0	-3	-1	-4	-4	0	-2	11	1
W	-2	-4	-4	-4	0	-2	0	-2	0	0	0	-3	-1	0	-4	4	-1	-3	-1	6	-6	4	-2	11	1
I	-3	-5	-5	-5	-5	-1	-6	-4	0	-4	3	0	-5	0	-4	-5	-4	-1	3	-4	-5	-3	-4	11	1
C	1	-3	-5	-5	-4	-1	-4	-4	0	0	3	0	-4	-4	-3	-1	1	-1	0	-5	-5	-3	-4	11	1
R	-3	-3	-5	-3	0	-4	3	-3	-2	1	1	2	0	1	0	4	-3	-3	-4	-4	-5	0	0	11	1
W	-3	-5	-4	-5	-4	2	-2	-3	-1	0	1	-2	-4	-5	-3	0	-2	0	0	8	-5	2	-4	11	1
M	-3	-5	-4	-5	-4	2	-2	-4	2	0	2	3	-1	-5	-3	-1	-2	0	1	2	-5	1	-4	11	1
40	-2	-5	-4	-5	-5	6	-5	0	0	-5	0	3	-5	-6	-4	-5	-4	-4	1	7	-6	4	-5	11	1
W																									

**Figure 1/5**





**Figure 4/5**

**Figure S/5**

	10	20	30	40
ATAT1	Y - - T Y E M L G I I H L T V V N G H R - P P P P S P G T L G N A L D P			
ATAT9	V - - N Y K L T G I K L V V N G H P - P P P P K P G Q P G H L V C N H R T V L			
ATAT7	F S G C R L T V T N D Y V N H N - D L I S A D R K R G C C L F V C T H R T L M			
ATAT8	A F S G I H L T L T I - V H N A - E R V R Q L A H D T N L P L V Y V C N H R T L L			
ATAT6	S - F F W A G V K I Q I V H K G Q L E M V K A A T E T N G L E L V Y V P C H R S H I D Y			
PLSB_ECOLI	I V D W W A G V K I Q I V H K G Q L E M V K A A T E T N G L E L V V C N H R S S D I D W			
PLSB_MOUSE	V - D W W A G V K I Q I V H K G Q L E M V K A A T E T N G L E L V I S N H R S S D I D W			
ATLPAAT1	V - D W W A G V K V Q L H A D E E T Y - - R L M G K E H A A L I I S N H R S S D I D W			
Jojoba AT	L - - W P F L F E K I N K T K V I F S G D K V P C E D R V L I A N H R T E V D W			
Maize AT	V T G R M L M W I L G N P I I T - - - I E G S E F S N T R A I Y I C N H A S S P I D A			
ATAT11	I I G G L V I W I Y G I I P I I K - - - I Q G S E S Y G N A I Y I - A N H Q S N Y D M I			
PLSC_COCO	F - G R L A P L - F G L L K K V - - - K P T D A E E N L A K K P A V Y V S N H Q S F L D I			
PLSC_LIM	F - Y H V M K L M L G L I D N V I E - - - G L E E N L P S S D T P A V Y V S N H Q S F L D I			
PLSC_ECOLI	W A S I S I Y P F Y K I N I E - - - I I N Q - K G E A A T E E P E R P G A I V S N H H V S Y L D I			
PLSC_YEAST	M D S N P K T T S T E - - - I I R R - K G K P A R - - R E I A P P T V V S N H H V S Y L D I			
ATAT2	R C I L F S F G - Y Q W - - - T G V V K - K G K P A R - - R E I A P P T V V S N H H V S Y L D I			
ATAT3	M I C S F F V A S - - - W - - - T G V V K - K G K P A R - - R E I A P P T V V S N H H V S Y L D I			
ATAT70				
ATAT4				

ATAT1  
ATAT9  
ATAT7  
ATAT8  
ATAT6  
PLSB\_ECOLI  
PLSB\_MOUSE  
ATLPAAT1  
Jojoba AT  
Maize AT  
ATAT11  
PLSC\_COCO  
PLSC\_LIM  
PLSC\_ECOLI  
PLSC\_YEAST  
ATAT2  
ATAT3  
ATAT10  
ATAT4

[illegible]

Figure 2  
2/3



220	230	240	250
ATAT71	VANYYVQKVI	IGAVLGFEC	TELTRKDKYLLLG
ATAT9	VEVYIQRV	VLGGTLGFEC	
ATAT7	VTANYYVQR	ILAAATLGFEC	
ATAT8	DGKLKFEVAN	NNVDIGKALLDFFE	
ATAT6	NGKVNFBVAN	HHVQHEIGN	
PLSB_ECOLI	LSKLRNLGQ	GYV-NFGEPP	MLPMT--YLNQH
PLSB_MOUSE	VIRMLRKNN	YGYVVRVDF	SLKE--YL-EG
ATLPAAT1	V--VHVH	IKRSMKDDL	LPESDDAIAQWC--RDQ
Jobba AT	V--VHVH	IKRSMKDDL	LPESDDAIAQWC--RDQ
Malze AT	V--VHVH	IKRSMKDDL	LPESDDAIAQWC--RDQ
ATAT11	E--VVIH	IRINLTQI	PNQEKDINAWL--MNTFF
PLSC_COCO	H--YVEM	IHALYVVDH	LPESQKPLV--S--KG
PLSC_LIM	D--YVKM	IHDYVVRN	LPASQKPLG--S--TNRSS
PLSC_ECOLI	E--LAAH	CRSIMEQK	IAELDKVEVARE--AAGK
PLSC_YEAST	E--FAEK	--VRDQ	MMVDTL--KEIIGYSP--AIND
ATAT2	V--LCNEA	RSKIAES	MLMATEGNLILSE--LGLS
ATAT3	--DDPK	LYASNVR	KLLMA--LGLS
ATAT10	G--ETG	IEFAERV	VRDMMISLRAGLKKKVP--WDGY
ATAT4			

Figure 2  
3/3

9/10

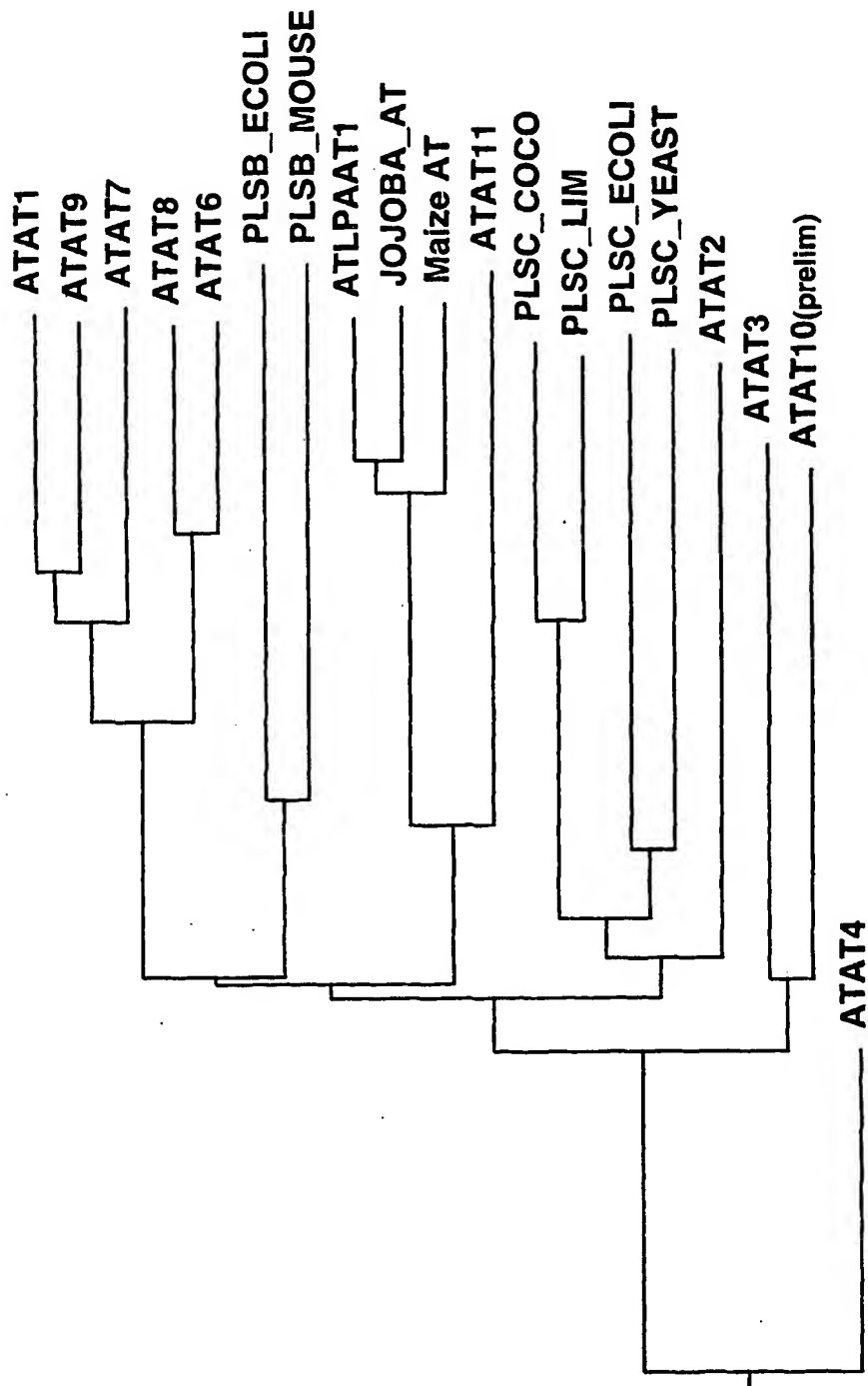


Figure 3 1/2

10/10

		Percent Similarity																			
Percent Divergence		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	
	1		67.0	58.2	47.2		11.6	11.5	12.0	10.5	12.4	13.4	12.9	12.8	12.3	12.3	11.8	16.3	14.4	14.4	
	2	31.4		63.9	44.8	42.8	12.9	12.4	14.9	14.4	13.9	16.5	11.3	12.4	12.4	11.3	11.2	13.4	13.7	14.4	
	3	40.2	35.8		44.8	44.8	12.9	14.4	14.9	13.4	11.3	12.9	12.4	12.9	11.9	13.9	13.4	13.4	17.1	14.4	
	4	49.7	50.0	50.3		67.2	10.8	13.3	11.8	11.8	10.8	16.4	11.8	11.8	12.8	13.3	12.3	17.4	15.1	12.8	
	5	50.3	50.3	48.6	25.7		12.3	12.3	13.8	12.8	12.3	12.3	12.3	12.8	12.8	12.3	12.8	15.9	13.7	15.9	
	6	85.6	86.3	85.6	86.2	86.1		28.5	12.6	12.1	11.6	9.7	13.9	14.3	14.8	11.8	17.6	13.5	12.3	10.6	
	7	83.8	86.8	82.8	82.7	84.3	66.2		12.6	13.9	12.9	13.1	12.4	13.3	14.3	13.8	15.0	11.7	11.6	10.0	
	8	82.9	78.4	81.2	83.1	81.2	83.6	85.1		82.3	78.0	31.6	12.4	12.8	13.3	15.8	13.9	12.2	16.4	14.4	
	9	83.5	77.8	81.8	85.9	84.6	85.6	87.1	18.2		77.5	32.1	11.9	14.3	13.3	16.3	15.5	12.4	15.1	12.0	
	10	83.5	82.4	84.1	87.6	85.1	84.4	87.1	22.5	22.5		30.6	13.9	16.7	12.8	16.3	14.4	12.9	16.4	12.0	
	11	84.1	81.4	83.1	85.1	85.1	90.3	87.4	65.4	65.0	67.5		14.4	14.8	11.8	12.8	13.4	14.6	15.8	12.9	
	12	83.6	84.6	84.6	83.3	83.2	84.6	86.8	80.9	82.2	81.1	82.4		66.7	27.9	28.4	23.5	14.9	17.1	12.9	
	13	82.1	84.0	81.7	81.0	82.0	87.6	86.9	79.5	80.5	79.5	81.6	33.3		26.1	28.1	19.3	14.8	14.4	15.3	
	14	83.2	81.4	85.0	83.6	82.9	89.1	86.8	82.1	81.6	80.5	84.3	71.3	71.6		30.0	19.3	18.7	17.8	12.3	
	15	83.1	80.6	83.0	79.3	81.0	88.2	87.4	81.4	82.0	79.8	84.1	70.3	70.1	62.8		20.3	15.3	15.1	13.8	
	16	82.7	82.6	83.9	78.4	78.8	80.8	86.1	82.9	82.2	81.6	84.5	72.3	75.0	76.5	73.5		18.2	15.8	17.6	
	17	83.7	82.0	86.6	78.4	80.2	86.7	89.8	86.4	85.5	85.5	84.4	80.1	78.2	78.6	80.6	77.8		30.8	17.2	
	18	78.5	82.5	82.5	81.7	81.8	88.7	87.1	79.1	80.5	78.9	82.8	81.8	78.1	76.1	78.1	79.3	64.8		18.5	
	19	84.7	84.8	84.8	84.7	85.5	86.5	83.6	87.4	86.5	87.6	91.0	85.0	83.1	85.7	81.8	83.7	79.4	74.1		
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	

ATAT1  
 ATAT9  
 ATAT7  
 ATAT8  
 ATAT6  
 PLSB\_ECOLI  
 PLSB\_MOUSE  
 ATLPAATI  
 JOJOBA\_AT  
 Maize AT  
 ATAT11  
 PLSC\_COCO  
 PLSC LIM  
 PLSC\_ECOLI  
 PLSC\_YEAST  
 ATAT2  
 ATAT3  
 ATAT10(prelim)  
 ATAT4

Figure 3 2/2

## SEQUENCE LISTING

<110> Lassner, Michael W  
Emig, Robin A  
Ruezinsky, Diane  
Van Eenennaam, Alison

<120> Novel Plant Acyltransferases

<130> 17029/00/WO

<140>

<141>

<150> 60/101,939

<151> 1998-09-25

<160> 241

<170> PatentIn Ver. 2.0

<210> 1

<211> 869

<212> DNA

<213> Arabidopsis sp.

<400> 1

```

atggttcacg cgaccaagtc agccacaacg attccaaaag aacgcttaaa gaaccgcata 60
gtcttccatg atgggcgttt agcgcaacgt ccaactccgt taaacgccat tatcacatac 120
ctatggcttc cttttggttt catctctcca tcattcgcgt ctacttcaac ctccctttac 180
ctgaaagatt tgtccgttac acttacgaga tgctcgggat ccacttaacc attcgtgggc 240
atcgtcctcc acctccttcc cctggaactc ttggcaacct ctatgtcctt aaccaccgta 300
ccgcgcttga tcccatcatc gtcgctattg ctcttgagc taagatctgt tgcgtcactt 360
acagtgtctc tcgtctctcc cttatgcttt ctctatttcc tgctgttgcc ctcacccgtg 420
accgtgccac cgatgctgcc aacatgagaa aacttctcga gaaaggcgac ttggtgatat 480
gtccggaagg cacgacgtgt agagaagagt atctactgag atttagcgct ctattcgcag 540
agctaagcga ccggattgtg ccagtagcga tgaactgtaa acaagggaatg ttcaacggga 600
ccacagttag ggggtgtgaag ttttgggacc cttacttctt cttcatgaac ccaagaccaa 660
gctatgaagc cactttcttg gatcgtttgc ctgaagaaat gactgtcaac ggtggtggca 720
agactcctat agaggtggct aattacgtcc agaaagtat cggcgcggtt ttgggcttcg 780
aatgcaccga acttactcgc aaggataaat atcttttgct tggaggtaat gacggcaagg 840
tggagtctat caacaacacc aagaagtga 869

```

<210> 2

<211> 289

<212> PRT

<213> Arabidopsis sp.

<400> 2

```

Met Val His Ala Thr Lys Ser Ala Thr Thr Ile Pro Lys Glu Arg Leu
 1             5             10             15

Lys Asn Arg Ile Val Phe His Asp Gly Arg Leu Ala Gln Arg Pro Thr
 20             25             30

Pro Leu Asn Ala Ile Ile Thr Tyr Leu Trp Leu Pro Phe Gly Phe Ile
 35             40             45

Leu Ser Ile Ile Arg Val Tyr Phe Asn Leu Pro Leu Pro Glu Arg Phe
 50             55             60

Val Arg Tyr Thr Tyr Glu Met Leu Gly Ile His Leu Thr Ile Arg Gly
 65             70             75             80

His Arg Pro Pro Pro Ser Pro Gly Thr Leu Gly Asn Leu Tyr Val
 85             90             95

Leu Asn His Arg Thr Ala Leu Asp Pro Ile Ile Val Ala Ile Ala Leu
100             105             110

```

## SEQUENCE LISTING

<110> Lassner, Michael W  
Emig, Robin A  
Ruezinsky, Diane  
Van Eenennaam, Alison

<120> Novel Plant Acyltransferases

<130> 17029/00/WO

<140>

<141>

<150> 60/101,939

<151> 1998-09-25

<160> 241

<170> PatentIn Ver. 2.0

<210> 1

<211> 869

<212> DNA

<213> Arabidopsis sp.

<400> 1

```

atggttcatg cgaccaagtc agccacaacg attccaaaag aacgcttaaa gaaccgcata 60
gtcttccatg atgggcggtt agcgcaacgt ccaactccgt taaacgccat tatcacatac 120
ctatggcttc cttttgggtt catctctcca tcattcgctt ctacttcaac ctccctttac 180
ctgaaagatt tgtccgttac acttacgaga tgctcgggat ccacttaacc attcgtgggc 240
atcgctctcc acctccttcc cctggaactc ttggcaacct ctatgtcctt aaccaccgta 300
ccgcgcttga tcccatcacc gtcgctattg ctcttggaag taagatctgt tgcgtcactt 360
acagtgtctc tcgtctctcc cttatgcttt ctctattccc tgctgttgcc ctccaccgtg 420
accgtgccac cgatgctgcc aacatgagaa aacttctcga gaaaggcgac ttggtgatat 480
gtccggaagg cagcagctgt agagaagagt atctactgag atttagcgct ctattcgag 540
agctaagcga ccggattgtg ccagtagcga tgaactgtaa acaagggaatg ttcaacggga 600
ccacagttag ggggtgtgaag ttttgggacc cttacttctt cttcatgaac ccaagaccaa 660
gctatgaagc cactttcttg gatcgtttgc ctgaagaaat gactgtcaac ggtggtggca 720
agactcctat agaggtggct aattacgtcc agaaagttat cggcgcggtt ttgggcttcg 780
aatgcaccga acttactcgc aaggataaat atcttttgct tggaggtaat gacggcaagg 840
tggagtctat caacaacacc aagaagtga                                     869

```

<210> 2

<211> 289

<212> PRT

<213> Arabidopsis sp.

<400> 2

```

Met Val His Ala Thr Lys Ser Ala Thr Thr Ile Pro Lys Glu Arg Leu
  1              5              10              15

Lys Asn Arg Ile Val Phe His Asp Gly Arg Leu Ala Gln Arg Pro Thr
      20              25              30

Pro Leu Asn Ala Ile Ile Thr Tyr Leu Trp Leu Pro Phe Gly Phe Ile
      35              40              45

Leu Ser Ile Ile Arg Val Tyr Phe Asn Leu Pro Leu Pro Glu Arg Phe
      50              55              60

Val Arg Tyr Thr Tyr Glu Met Leu Gly Ile His Leu Thr Ile Arg Gly
      65              70              75              80

His Arg Pro Pro Pro Pro Ser Pro Gly Thr Leu Gly Asn Leu Tyr Val
      85              90              95

Leu Asn His Arg Thr Ala Leu Asp Pro Ile Ile Val Ala Ile Ala Leu
      100             105             110

```

Gly Arg Lys Ile Cys Cys Val Thr Tyr Ser Val Ser Arg Leu Ser Leu  
 115 120 125

Met Leu Ser Pro Ile Pro Ala Val Ala Leu Thr Arg Asp Arg Ala Thr  
 130 135 140

Asp Ala Ala Asn Met Arg Lys Leu Leu Glu Lys Gly Asp Leu Val Ile  
 145 150 155 160

Cys Pro Glu Gly Thr Thr Cys Arg Glu Glu Tyr Leu Leu Arg Phe Ser  
 165 170 175

Ala Leu Phe Ala Glu Leu Ser Asp Arg Ile Val Pro Val Ala Met Asn  
 180 185 190

Cys Lys Gln Gly Met Phe Asn Gly Thr Thr Val Arg Gly Val Lys Phe  
 195 200 205

Trp Asp Pro Tyr Phe Phe Phe Met Asn Pro Arg Pro Ser Tyr Glu Ala  
 210 215 220

Thr Phe Leu Asp Arg Leu Pro Glu Glu Met Thr Val Asn Gly Gly Gly  
 225 230 235 240

Lys Thr Pro Ile Glu Val Ala Asn Tyr Val Gln Lys Val Ile Gly Ala  
 245 250 255

Val Leu Gly Phe Glu Cys Thr Glu Leu Thr Arg Lys Asp Lys Tyr Leu  
 260 265 270

Leu Leu Gly Gly Asn Asp Gly Lys Val Glu Ser Ile Asn Asn Thr Lys  
 275 280 285

Lys

<210> 3  
 <211> 939  
 <212> DNA  
 <213> Arabidopsis sp.

<400> 3  
 atgacgagct ttactacttc ccttcatgct gtcccgagtg aaaaatttat gggcgaaaca 60  
 agacgtactg gcattcaatg gtctaaccgc tctttaagac atgacacctta cagatttctt 120  
 gataagaaat cacctagatc aagtcaattg gcaagagata tcaactgtgag agcagatctt 180  
 tcaggagctg caaccctga ctcttctttt cctgaaccag agattaagtt gagctcaaga 240  
 ctacagagga tattcttttg tgttgttgct ggcatttcgg ctacttttct cattgtcctg 300  
 atgattattg ggcacccgtt cgtccttctc ttcgatccct ataggagaaa attccaccac 360  
 ttcatgtgta aactttgggc ttccataagc atttatccgt tttacaaaat caacatcgag 420  
 gggttggaga atctgccatc atcagacact cctgctgtat atgtttcaaa ccaccaaagt 480  
 tttctggata tctacacact tcttagtctt ggaaaaagct ttaagttcat cagcaagaca 540  
 gggatattcg taattcccat catcggttgg gccatgtcca tgatgggtgt cgttcccttg 600  
 aagcggatgg acccaagaag ccaagtggat tgcttaaaac gctgcatgga acttttaaag 660  
 aagggagcat ctgtgttttt ctcccagaa ggaacacgga gtaaggatgg tcggttaggt 720  
 tctttcaaga aaggcgcatc tacagtggct gcgaagaccg gagttgcagt agttccaata 780  
 acgctaattg gaacaggcaa aatcatgcca acgggtagtg aaggtatact gaaccatggg 840  
 aatgtgagag ttatcatcca taaaccaata catggaagca aagcggatgt tctttgcaac 900  
 gaggccagaa gcaagattgc agaatcaatg gatctctaa 939

<210> 4  
 <211> 312  
 <212> PRT  
 <213> Arabidopsis sp.

<400> 4  
 Met Thr Ser Phe Thr Thr Ser Leu His Ala Val Pro Ser Glu Lys Phe  
 1 5 10 15

Met Gly Glu Thr Arg Arg Thr Gly Ile Gln Trp Ser Asn Arg Ser Leu

20					25					30					
Arg	His	Asp	Pro	Tyr	Arg	Phe	Leu	Asp	Lys	Lys	Ser	Pro	Arg	Ser	Ser
		35					40					45			
Gln	Leu	Ala	Arg	Asp	Ile	Thr	Val	Arg	Ala	Asp	Leu	Ser	Gly	Ala	Ala
	50					55					60				
Thr	Pro	Asp	Ser	Ser	Phe	Pro	Glu	Pro	Glu	Ile	Lys	Leu	Ser	Ser	Arg
	65					70					75				80
Leu	Arg	Gly	Ile	Phe	Phe	Cys	Val	Val	Ala	Gly	Ile	Ser	Ala	Thr	Phe
				85					90					95	
Leu	Ile	Val	Leu	Met	Ile	Ile	Gly	His	Pro	Phe	Val	Leu	Leu	Phe	Asp
			100					105					110		
Pro	Tyr	Arg	Arg	Lys	Phe	His	His	Phe	Ile	Ala	Lys	Leu	Trp	Ala	Ser
			115				120					125			
Ile	Ser	Ile	Tyr	Pro	Phe	Tyr	Lys	Ile	Asn	Ile	Glu	Gly	Leu	Glu	Asn
	130					135					140				
Leu	Pro	Ser	Ser	Asp	Thr	Pro	Ala	Val	Tyr	Val	Ser	Asn	His	Gln	Ser
	145					150					155				160
Phe	Leu	Asp	Ile	Tyr	Thr	Leu	Leu	Ser	Leu	Gly	Lys	Ser	Phe	Lys	Phe
				165					170					175	
Ile	Ser	Lys	Thr	Gly	Ile	Phe	Val	Ile	Pro	Ile	Ile	Gly	Trp	Ala	Met
			180					185					190		
Ser	Met	Met	Gly	Val	Val	Pro	Leu	Lys	Arg	Met	Asp	Pro	Arg	Ser	Gln
		195					200					205			
Val	Asp	Cys	Leu	Lys	Arg	Cys	Met	Glu	Leu	Leu	Lys	Lys	Gly	Ala	Ser
	210					215					220				
Val	Phe	Phe	Phe	Pro	Glu	Gly	Thr	Arg	Ser	Lys	Asp	Gly	Arg	Leu	Gly
	225					230					235				240
Ser	Phe	Lys	Lys	Gly	Ala	Phe	Thr	Val	Ala	Ala	Lys	Thr	Gly	Val	Ala
				245					250					255	
Val	Val	Pro	Ile	Thr	Leu	Met	Gly	Thr	Gly	Lys	Ile	Met	Pro	Thr	Gly
			260					265					270		
Ser	Glu	Gly	Ile	Leu	Asn	His	Gly	Asn	Val	Arg	Val	Ile	Ile	His	Lys
		275					280					285			
Pro	Ile	His	Gly	Ser	Lys	Ala	Asp	Val	Leu	Cys	Asn	Glu	Ala	Arg	Ser
	290					295					300				
Lys	Ile	Ala	Glu	Ser	Met	Asp	Leu								
	305					310									

&lt;210&gt; 5

&lt;211&gt; 1197

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis sp.

&lt;400&gt; 5

```

atggaatcag agctcaaaga tttgaattcg aattcgaatc ctccgtcgag caaagaggac 60
cggccgttac tgaatcaga atccgatttg gcggctgcca ttgaagagtt agacaaaaag 120
ttcgcacctt acgcgaggac cgatttgtat gggacgatgg gtttgggtcc tttcccgatg 180
acggagaata ttaaattggc gttgcattg gtgactcttg ttccattgcg gtttcttctc 240
tcgatgagca tcttgcttct ctattacttg attttagggg tatttacgct gttttctgct 300
ccttatcggt ggccagagga agaggaagat gaaggtggag ttgtttttca ggaagattat 360
gtcacatggt aaggttggaa acggactgtt atcgctccgt ctgggaggtt tctctctagg 420

```

```

gtttttgcttt tcgttttttg gtttttattgg attcacgaga gctgtccaga tcgagattca 480
gacatggatt ctaatcctaa aactacttct acagagatta accagaaagg ggaagccgcc 540
acggaggaac ctgaaagacc tggagccatt gtgtccaatc atgtttcgta cttggacatt 600
ttgtatcata tgtctgcttc ttttccaagt ttgtgttgcca agagatcagt gggcaaaactt 660
cctcttggtt gcctcattag caaatgcctt ggttgtgtct atgttcaaag agaagcaaaa 720
tcgcctgatt tcaaggggtg atctggcaca gtaaatgaaa gagttcgaga agctcatagc 780
aataaatctg ctccaactat tatgcttttt ccagaaggaa caactaccaa tggagactac 840
ttacttacat tcaagacagg tgcatttttt gctggaactc cagttcttcc ggtaatatata 900
aaatatccgt atgagcgctt cagtgtggca tgggatacca tatccggggc acgccacatt 960
ttattccttc tctgtcaagt cgtaaatacac ttggaagtca tacggttacc tgtatactac
1020
ccatcccaag aagagaaaga cgatcccaaa ctttatgcta gcaatgttcg gaaattaatg
1080
gccaccgagg gtaacttgat tctatcggag ttgggactta gcgacaaaag gatatatcac
1140
gcaactctca atggtaatct tagtcaaacc cgtgatttcc atcagaaaga agaataga
1197

```

<210> 6  
 <211> 398  
 <212> PRT  
 <213> Arabidopsis sp.

```

<400> 6
Met Glu Ser Glu Leu Lys Asp Leu Asn Ser Asn Ser Asn Pro Pro Ser
  1          5          10          15
Ser Lys Glu Asp Arg Pro Leu Leu Lys Ser Glu Ser Asp Leu Ala Ala
          20          25          30
Ala Ile Glu Glu Leu Asp Lys Lys Phe Ala Pro Tyr Ala Arg Thr Asp
          35          40          45
Leu Tyr Gly Thr Met Gly Leu Gly Pro Phe Pro Met Thr Glu Asn Ile
          50          55          60
Lys Leu Ala Val Ala Leu Val Thr Leu Val Pro Leu Arg Phe Leu Leu
          65          70          75          80
Ser Met Ser Ile Leu Leu Leu Tyr Tyr Leu Ile Cys Arg Val Phe Thr
          85          90          95
Leu Phe Ser Ala Pro Tyr Arg Gly Pro Glu Glu Glu Glu Asp Glu Gly
          100          105          110
Gly Val Val Phe Gln Glu Asp Tyr Ala His Met Glu Gly Trp Lys Arg
          115          120          125
Thr Val Ile Val Arg Ser Gly Arg Phe Leu Ser Arg Val Leu Leu Phe
          130          135          140
Val Phe Gly Phe Tyr Trp Ile His Glu Ser Cys Pro Asp Arg Asp Ser
          145          150          155          160
Asp Met Asp Ser Asn Pro Lys Thr Thr Ser Thr Glu Ile Asn Gln Lys
          165          170          175
Gly Glu Ala Ala Thr Glu Glu Pro Glu Arg Pro Gly Ala Ile Val Ser
          180          185          190
Asn His Val Ser Tyr Leu Asp Ile Leu Tyr His Met Ser Ala Ser Phe
          195          200          205
Pro Ser Phe Val Ala Lys Arg Ser Val Gly Lys Leu Pro Leu Val Gly
          210          215          220
Leu Ile Ser Lys Cys Leu Gly Cys Val Tyr Val Gln Arg Glu Ala Lys
          225          230          235          240
Ser Pro Asp Phe Lys Gly Val Ser Gly Thr Val Asn Glu Arg Val Arg

```



	245		250		255
Glu Ala His	Ser Asn Lys Ser Ala Pro Thr Ile Met Leu Phe Pro Glu				
	260		265		270
Gly Thr Thr	Thr Asn Gly Asp Tyr Leu Leu Thr Phe Lys Thr Gly Ala				
	275		280		285
Phe Leu Ala Gly Thr Pro Val Leu Pro Val Ile Leu Lys Tyr Pro Tyr					
	290		295		300
Glu Arg Phe Ser Val Ala Trp Asp Thr Ile Ser Gly Ala Arg His Ile					
305		310		315	320
Leu Phe Leu Leu Cys Gln Val Val Asn His Leu Glu Val Ile Arg Leu					
	325		330		335
Pro Val Tyr Tyr Pro Ser Gln Glu Glu Lys Asp Asp Pro Lys Leu Tyr					
	340		345		350
Ala Ser Asn Val Arg Lys Leu Met Ala Thr Glu Gly Asn Leu Ile Leu					
	355		360		365
Ser Glu Leu Gly Leu Ser Asp Lys Arg Ile Tyr His Ala Thr Leu Asn					
	370		375		380
Gly Asn Leu Ser Gln Thr Arg Asp Phe His Gln Lys Glu Glu					
385		390		395	

&lt;210&gt; 7

&lt;211&gt; 1131

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis sp.

&lt;400&gt; 7

```

atgagcagta cggcagggag gctcgtgact tcaaaatccg agcttgacct cgatcacct 60
aacatcgaag attaccttcc ttctggttct tccatcaatg aacctcgcgg caagctcagc 120
ctgcgtgatt tgctagacat ctctccaacg ctactgaag ctgctgggtgc cattgttgat 180
gactcgttca caagatgttt caaatcaaat cctccagaac cttggaactg gaatatttac 240
ttattcccac tatactgctt tggggttgtt gttagatact gtatcctctt tcccttgagg 300
tgcttcaact tagcttttgg gtggattatt ttcctttcat tgtttatccc tgtaaatgcg 360
ttgctgaaag gtcaagatag gttgaggaaa aagatagaga gggctcttggg ggaaatgatt 420
tgcagctttt ttgtcgctc atggaccgga gttgtcaa atcacgggccc acgtcctagc 480
atccgtccta agcaggtcta tgttgccaac catacttcaa tgattgattt catcgtattg 540
gagcagatga ccgcatttgc tgttataatg cagaagcatc ctgggtgggt tggctctctg 600
caaagcacaa tattagagag tgtgggatgt atctgggtca atcggtcaga ggcaaggat 660
cgtgaaattg tagcaaaaaa gtttaaggac catgtccaag gagctgacag taatcctctt 720
ctcatatttc ccgaagggac atgtgtaaat aataattaca cagtgatgtt taagaagggt 780
gcttttgaat tggactgcac tgtttgtcca attgcaatta aatacaacaa gatttttgtt 840
gacgccttct ggaatagcag aaaacaatca tttactatgc acttgctgca actcatgaca 900
tcatgggctg ttgtatgtga agtgtgttac ttggaaccac aaaccataag gcccggtgaa 960
acaggaattg aatttgcaga gagggtcaga gacatgatat ctcttcgggc gggctctcaa
1020
aagggtccctt gggatggata cttgaagtat tcgagaccaa gcccgaagca tagtgaacgc
1080
aagcaacaga gtttcgcaga gtcgatcctg gctagattgg aagagaagtg a
1131

```

&lt;210&gt; 8

&lt;211&gt; 376

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis sp.

&lt;400&gt; 8

Met Ser Ser Thr Ala Gly Arg Leu Val Thr Ser Lys Ser Glu Leu Asp					
1		5		10	15
Leu Asp His Pro Asn Ile Glu Asp Tyr Leu Pro Ser Gly Ser Ser Ile					
	20		25		30

Asn Glu Pro Arg Gly Lys Leu Ser Leu Arg Asp Leu Leu Asp Ile Ser  
                   35                  40                  45  
 Pro Thr Leu Thr Glu Ala Ala Gly Ala Ile Val Asp Asp Ser Phe Thr  
                   50                  55                  60  
 Arg Cys Phe Lys Ser Asn Pro Pro Glu Pro Trp Asn Trp Asn Ile Tyr  
                   65                  70                  75                  80  
 Leu Phe Pro Leu Tyr Cys Phe Gly Val Val Val Arg Tyr Cys Ile Leu  
                   85                  90                  95  
 Phe Pro Leu Arg Cys Phe Thr Leu Ala Phe Gly Trp Ile Ile Phe Leu  
                   100                  105                  110  
 Ser Leu Phe Ile Pro Val Asn Ala Leu Leu Lys Gly Gln Asp Arg Leu  
                   115                  120                  125  
 Arg Lys Lys Ile Glu Arg Val Leu Val Glu Met Ile Cys Ser Phe Phe  
                   130                  135                  140  
 Val Ala Ser Trp Thr Gly Val Val Lys Tyr His Gly Pro Arg Pro Ser  
                   145                  150                  155                  160  
 Ile Arg Pro Lys Gln Val Tyr Val Ala Asn His Thr Ser Met Ile Asp  
                   165                  170                  175  
 Phe Ile Val Leu Glu Gln Met Thr Ala Phe Ala Val Ile Met Gln Lys  
                   180                  185                  190  
 His Pro Gly Trp Val Gly Leu Leu Gln Ser Thr Ile Leu Glu Ser Val  
                   195                  200                  205  
 Gly Cys Ile Trp Phe Asn Arg Ser Glu Ala Lys Asp Arg Glu Ile Val  
                   210                  215                  220  
 Ala Lys Lys Leu Arg Asp His Val Gln Gly Ala Asp Ser Asn Pro Leu  
                   225                  230                  235                  240  
 Leu Ile Phe Pro Glu Gly Thr Cys Val Asn Asn Asn Tyr Thr Val Met  
                   245                  250                  255  
 Phe Lys Lys Gly Ala Phe Glu Leu Asp Cys Thr Val Cys Pro Ile Ala  
                   260                  265                  270  
 Ile Lys Tyr Asn Lys Ile Phe Val Asp Ala Phe Trp Asn Ser Arg Lys  
                   275                  280                  285  
 Gln Ser Phe Thr Met His Leu Leu Gln Leu Met Thr Ser Trp Ala Val  
                   290                  295                  300  
 Val Cys Glu Val Trp Tyr Leu Glu Pro Gln Thr Ile Arg Pro Gly Glu  
                   305                  310                  315                  320  
 Thr Gly Ile Glu Phe Ala Glu Arg Val Arg Asp Met Ile Ser Leu Arg  
                   325                  330                  335  
 Ala Gly Leu Lys Lys Val Pro Trp Asp Gly Tyr Leu Lys Tyr Ser Arg  
                   340                  345                  350  
 Pro Ser Pro Lys His Ser Glu Arg Lys Gln Gln Ser Phe Ala Glu Ser  
                   355                  360                  365  
 Ile Leu Ala Arg Leu Glu Glu Lys  
                   370                  375

&lt;210&gt; 9

&lt;211&gt; 965

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis sp.

&lt;400&gt; 9

```

gttgtaagt tacaagtctc ttcaaaaaca cacacacacg tctctcttca cagccaatca 60
tcgatcacag ctcgattttc ctttattgtt ccgttggttt tcttgagnat ttttctttct 120
tgggatcatc aaactngtcg gtaaggwaac ttcacggacg gatcttcaat gttgagctgt 180
tctaattgga ccgtcgtgat cgcaaccgcc atgggttgct caagcaccgc tctgtttctc 240
gccatggctc gtcaattcca tggaaatcat caaaatccta aggttcttga tcagactcta 300
cgacccattc tccgttcttg tctatcttca gaggaacga agaaacaggg gaagaagata 360
aagaaagtgc gggtcgcgga taatgtgaaa gatacgaag gtaacgggga agagtaccgg 420
aggaggggaat tgaaccggaa aagcgtaccg aagccagtga ctaaaccggg aaagaccggg 480
tctatgtgta gaatctctac catgccagcg aaccggatgg ctctgtacaa tgggattctt 540
agagaccgag atcacagagt tcaatattct tattgacttt ttcttcttga ttagtcaata 600
gatttaggtt ttgtaaatct ttcttttggt tttcggtaat attagatttt ttcttggaaa 660
tttcagatat tgtagacttt gtagattggg ggtctctctt ttctcccttt ttgtgtctca 720
tagtagtagg tgggtttctt atgctccact tatctactta cttgttttaa atcaagtgat 780
gatgtaaata attgacatgt aagtagtcat tagaaatttg aaaaggcaaa tgaaagaata 840
taaatttgta aaaacatagt gtgcctattg tacatataaa ctctcttttg ttggggatat 900
ctatggaatt tatattgatt gtgttgaaaa aacaaaaaaa aaaaaaaaaa aaaaaaaaaa 965
aaaaa

```

&lt;210&gt; 10

&lt;211&gt; 1593

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis sp.

&lt;400&gt; 10

```

atgtccggtg ataagatctc gactcttcaa gctcttgtct tcttcttgta ccggtttttc 60
attctccgtc gttgggtgca tcttagccct aaacaaaaat accaaaaatg cccttctcac 120
ggcctccacc aatatcaaga cctatcgaat cacactttga tattcaacgt cgaaggagct 180
ctactcaaat caaactcttt attccttac tctatgggtg tggcattcga agccggaggg 240
gtgataaggt cacttttctt cttagtctct tatccattta taagcttgat gagctacgaa 300
atgggcttga agacgatggt gatgctgagc ttctttggag ttaaaaagga aagcttccga 360
gtggggaaat cagttttgcc taagtatttt ctagaagatg ttgggctcga gatgttccag 420
gttttgaaaa gaggagggca gagagtgtct gtgagtgtat taccacaagt tatgattgat 480
gtattcttgc gagattactt ggagatagaa gttgtggctg gaagagacat gaaaatggct 540
gggtggtact acctaggcat cgtggaggat aagaagaacc ttgaaattgc ttttgataaa 600
gtggttcaag aagaaagact tggtagtggt cgtcgtctta ttggcatcac ttctttaa 660
tcgccaagtc acagatctct cttctctcaa ttttggcagg aaatttactt cgtcgaat 720
tcagacaaga aaagtggca aaccctacca caagatcaat accctaaacc attgattttc 780
cacgatggtc gtttagccgt taagccaaca cctttaaaca cactcgtatt attcatgtg 840
gccccattcg ccgcgctctt agccgctgca agactcgtct tcggcctaaa cttaccttac 900
tccctagcca atcccttctt cgccttttcc ggtatccacc ttactctcac cgtcaacaac 960
cacaacgacc taatatccgc cgacagaaaa agaggttgct tctttgtgtg taaccataga
1020
acgttattgg acccacttta catttcatac gctctaagaa agaaaaacat gaaagccgtg
1080
acgtatagtc taagcagatt atctgagctt ctggctccga tcaagaccgt tagattgact
1140
cgtgatcgag tcaaagatgg tcaagccatg gagaaattgc tgagccaggg agatctcgtg
1200
gtttgtccgg aagggactac gtgtagagag ctttacttgc ttgggtttag tccacttttc
1260
tctgaggttt gtgacgtcat cgtacctgtt gctattgact cacacgtgac tttcttctat
1320
ggcacgacgg ctagtggctt taaggcattt gatccattt tcttctttt gaatcctttc
1380
ccttcttaca ccgtcaaatt gcttgaccct gtctctggaa gtagctcgtc cacgtgtcga
1440
ggagtccctg acaatggaaa agttaacttc gaggtggcta atcacgtgca gcatgagatc
1500
gggaatgcct tggggtttga gtgcaccaac ctcacgagaa gagataagta cttgatcttg
1560
gccggttaata acggagttgt caagaaaaaa taa
1593

```

&lt;210&gt; 11

&lt;211&gt; 530

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis sp.

&lt;400&gt; 11

```

Met Ser Gly Asn Lys Ile Ser Thr Leu Gln Ala Leu Val Phe Phe Leu
 1      5      10      15
Tyr Arg Phe Phe Ile Leu Arg Arg Trp Cys His Arg Ser Pro Lys Gln
      20      25      30
Lys Tyr Gln Lys Cys Pro Ser His Gly Leu His Gln Tyr Gln Asp Leu
      35      40      45
Ser Asn His Thr Leu Ile Phe Asn Val Glu Gly Ala Leu Leu Lys Ser
      50      55      60
Asn Ser Leu Phe Pro Tyr Phe Met Val Val Ala Phe Glu Ala Gly Gly
      65      70      75      80
Val Ile Arg Ser Leu Phe Leu Leu Val Leu Tyr Pro Phe Ile Ser Leu
      85      90      95
Met Ser Tyr Glu Met Gly Leu Lys Thr Met Val Met Leu Ser Phe Phe
      100      105      110
Gly Val Lys Lys Glu Ser Phe Arg Val Gly Lys Ser Val Leu Pro Lys
      115      120      125
Tyr Phe Leu Glu Asp Val Gly Leu Glu Met Phe Gln Val Leu Lys Arg
      130      135      140
Gly Gly Lys Arg Val Ala Val Ser Asp Leu Pro Gln Val Met Ile Asp
      145      150      155      160
Val Phe Leu Arg Asp Tyr Leu Glu Ile Glu Val Val Val Gly Arg Asp
      165      170      175
Met Lys Met Val Gly Gly Tyr Tyr Leu Gly Ile Val Glu Asp Lys Lys
      180      185      190
Asn Leu Glu Ile Ala Phe Asp Lys Val Val Gln Glu Glu Arg Leu Gly
      195      200      205
Ser Gly Arg Arg Leu Ile Gly Ile Thr Ser Phe Asn Ser Pro Ser His
      210      215      220
Arg Ser Leu Phe Ser Gln Phe Cys Gln Glu Ile Tyr Phe Val Arg Asn
      225      230      235      240
Ser Asp Lys Lys Ser Trp Gln Thr Leu Pro Gln Asp Gln Tyr Pro Lys
      245      250      255
Pro Leu Ile Phe His Asp Gly Arg Leu Ala Val Lys Pro Thr Pro Leu
      260      265      270
Asn Thr Leu Val Leu Phe Met Trp Ala Pro Phe Ala Ala Val Leu Ala
      275      280      285
Ala Ala Arg Leu Val Phe Gly Leu Asn Leu Pro Tyr Ser Leu Ala Asn
      290      295      300
Pro Phe Leu Ala Phe Ser Gly Ile His Leu Thr Leu Thr Val Asn Asn
      305      310      315      320
His Asn Asp Leu Ile Ser Ala Asp Arg Lys Arg Gly Cys Leu Phe Val
      325      330      335
Cys Asn His Arg Thr Leu Leu Asp Pro Leu Tyr Ile Ser Tyr Ala Leu
      340      345      350
Arg Lys Lys Asn Met Lys Ala Val Thr Tyr Ser Leu Ser Arg Leu Ser

```

355 360 365  
 Glu Leu Leu Ala Pro Ile Lys Thr Val Arg Leu Thr Arg Asp Arg Val  
 370 375 380  
 Lys Asp Gly Gln Ala Met Glu Lys Leu Leu Ser Gln Gly Asp Leu Val  
 385 390 395 400  
 Val Cys Pro Glu Gly Thr Thr Cys Arg Glu Pro Tyr Leu Leu Arg Phe  
 405 410 415  
 Ser Pro Leu Phe Ser Glu Val Cys Asp Val Ile Val Pro Val Ala Ile  
 420 425 430  
 Asp Ser His Val Thr Phe Phe Tyr Gly Thr Thr Ala Ser Gly Leu Lys  
 435 440 445  
 Ala Phe Asp Pro Ile Phe Phe Leu Leu Asn Pro Phe Pro Ser Tyr Thr  
 450 455 460  
 Val Lys Leu Leu Asp Pro Val Ser Gly Ser Ser Ser Ser Thr Cys Arg  
 465 470 475 480  
 Gly Val Pro Asp Asn Gly Lys Val Asn Phe Glu Val Ala Asn His Val  
 485 490 495  
 Gln His Glu Ile Gly Asn Ala Leu Gly Phe Glu Cys Thr Asn Leu Thr  
 500 505 510  
 Arg Arg Asp Lys Tyr Leu Ile Leu Ala Gly Asn Asn Gly Val Val Lys  
 515 520 525  
 Lys Lys  
 530

<210> 12  
 <211> 1509  
 <212> DNA  
 <213> Arabidopsis sp.

<400> 12  
 atgggtatgg agcaagctgg aacgacatcg tattcgggtcg tgtcagagtt tgaaggaaca 60  
 atactgaaga acgcagattc attctcttac ttcattgctc tagccttcga agcagctggg 120  
 ctaattcggt tgcctatctt gttgtttcta tggcccgtaa tcacactcct tgacgttttc 180  
 agctacaaaa acgcagctct caagctcaag attttttag ccactgttgg tctacgtgaa 240  
 cggagatcg aatcagtggt tagagccgtt ctgccaaaat tctacatgga cgacgtaagc 300  
 atggacacgt ggagggtttt cagctcgtgt aagaagaggg tctgtgtcac gagaatgcct 360  
 cgagttatgg tggagaggtt tgctaaggag catcttagag cagatgaggt catcggtagc 420  
 gaactgattg taaaccggtt cggttttgtc accggtttga ttcgcgaaac ggatgttgat 480  
 cagtcgtctt tgaaccgtgt cgctaatttg tttgttggtc ggaggcctca actaggtctt 540  
 ggaaaaccgg ctttgaccgc ctctacaaat ttcttatcgt tatgtgagga gcatattcat 600  
 gcaccaatcc cggagaacta caaccacggt gaccaacaac ttcagctacy tccacttccg 660  
 gtgatatttc acgacggaag actagtgaag cggccaacgc cggccaccgc tctcatcatc 720  
 ctccatttgg tcccatttgg aatcattctc gccgtgatcc ggatctttct tggagccgtc 780  
 ctccatttgt gggccacacc ttacgtctct cagatattcg gtggccatat catcgtcaaa 840  
 ggaaagcctc ctcagccacc ggcggctgga aaatccggcg tgctctttgt gtgtactcac 900  
 agaaccctaa tggaccctgt ggtattatct tatgtcctcg gacgtagcat cccagccgtt 960  
 acttactcaa tctcgcgctt atcagagatc ttatctccca ttccaaccgt ccgattgaca  
 1020  
 agaatccgag atgtggatgc ggctaagatc aaacaacaac tgtcaaaagg agatctagt  
 1080  
 gtttgcctg agggaaccac ttgtcgtgaa cggtttttgt taagattcag cgcgcttttc  
 1140  
 gctgagttaa cggataggat tgttccggtt gcgatgaact acagagtcgg attcttccac  
 1200  
 gcgactacag cgagaggctg gaagggtttg gaccaattt tcttcttcat gaaccaaga  
 1260  
 cgggtttacg agattacgtt cttgaaccag cttcctatgg aggcaacatg ttcgtccggg  
 1320

aagagcccg c atgacgtggc gaactatgtt cagagaatct tggcgggtac gttagggttt  
 1380  
 gagtgaccca acttcacaag aaaagataag tatagggttc tcgctggaaa cgatgggaaca  
 1440  
 gtgtcgtact tgctgttgct agaccaattg aagaagggtg ttagcacttt cgagccttgt  
 1500  
 ctccattga  
 1509

<210> 13  
 <211> 502  
 <212> PRT  
 <213> Arabidopsis sp.

<400> 13  
 Met Val Met Glu Gln Ala Gly Thr Thr Ser Tyr Ser Val Val Ser Glu  
           1                  5                  10                  15  
 Phe Glu Gly Thr Ile Leu Lys Asn Ala Asp Ser Phe Ser Tyr Phe Met  
                   20                  25                  30  
 Leu Val Ala Phe Glu Ala Ala Gly Leu Ile Arg Phe Ala Ile Leu Leu  
                   35                  40                  45  
 Phe Leu Trp Pro Val Ile Thr Leu Leu Asp Val Phe Ser Tyr Lys Asn  
           50                  55                  60  
 Ala Ala Leu Lys Leu Lys Ile Phe Val Ala Thr Val Gly Leu Arg Glu  
           65                  70                  75                  80  
 Pro Glu Ile Glu Ser Val Ala Arg Ala Val Leu Pro Lys Phe Tyr Met  
                   85                  90                  95  
 Asp Asp Val Ser Met Asp Thr Trp Arg Val Phe Ser Ser Cys Lys Lys  
                   100                  105                  110  
 Arg Val Val Val Thr Arg Met Pro Arg Val Met Val Glu Arg Phe Ala  
                   115                  120                  125  
 Lys Glu His Leu Arg Ala Asp Glu Val Ile Gly Thr Glu Leu Ile Val  
           130                  135                  140  
 Asn Arg Phe Gly Phe Val Thr Gly Leu Ile Arg Glu Thr Asp Val Asp  
           145                  150                  155                  160  
 Gln Ser Ala Leu Asn Arg Val Ala Asn Leu Phe Val Gly Arg Arg Pro  
                   165                  170                  175  
 Gln Leu Gly Leu Gly Lys Pro Ala Leu Thr Ala Ser Thr Asn Phe Leu  
                   180                  185                  190  
 Ser Leu Cys Glu Glu His Ile His Ala Pro Ile Pro Glu Asn Tyr Asn  
           195                  200                  205  
 His Gly Asp Gln Gln Leu Gln Leu Arg Pro Leu Pro Val Ile Phe His  
           210                  215                  220  
 Asp Gly Arg Leu Val Lys Arg Pro Thr Pro Ala Thr Ala Leu Ile Ile  
           225                  230                  235                  240  
 Leu Leu Trp Ile Pro Phe Gly Ile Ile Leu Ala Val Ile Arg Ile Phe  
                   245                  250                  255  
 Leu Gly Ala Val Leu Pro Leu Trp Ala Thr Pro Tyr Val Ser Gln Ile  
                   260                  265                  270  
 Phe Gly Gly His Ile Ile Val Lys Gly Lys Pro Pro Gln Pro Pro Ala  
           275                  280                  285  
 Ala Gly Lys Ser Gly Val Leu Phe Val Cys Thr His Arg Thr Leu Met

290

295

300

Asp Pro Val Val Leu Ser Tyr Val Leu Gly Arg Ser Ile Pro Ala Val  
 305 310 315 320  
 Thr Tyr Ser Ile Ser Arg Leu Ser Glu Ile Leu Ser Pro Ile Pro Thr  
 325 330 335  
 Val Arg Leu Thr Arg Ile Arg Asp Val Asp Ala Ala Lys Ile Lys Gln  
 340 345 350  
 Gln Leu Ser Lys Gly Asp Leu Val Val Cys Pro Glu Gly Thr Thr Cys  
 355 360 365  
 Arg Glu Pro Phe Leu Leu Arg Phe Ser Ala Leu Phe Ala Glu Leu Thr  
 370 375 380  
 Asp Arg Ile Val Pro Val Ala Met Asn Tyr Arg Val Gly Phe Phe His  
 385 390 395 400  
 Ala Thr Thr Ala Arg Gly Trp Lys Gly Leu Asp Pro Ile Phe Phe Phe  
 405 410 415  
 Met Asn Pro Arg Pro Val Tyr Glu Ile Thr Phe Leu Asn Gln Leu Pro  
 420 425 430  
 Met Glu Ala Thr Cys Ser Ser Gly Lys Ser Pro His Asp Val Ala Asn  
 435 440 445  
 Tyr Val Gln Arg Ile Leu Ala Ala Thr Leu Gly Phe Glu Cys Thr Asn  
 450 455 460  
 Phe Thr Arg Lys Asp Lys Tyr Arg Val Leu Ala Gly Asn Asp Gly Thr  
 465 470 475 480  
 Val Ser Tyr Leu Ser Leu Leu Asp Gln Leu Lys Lys Val Val Ser Thr  
 485 490 495  
 Phe Glu Pro Cys Leu His  
 500

&lt;210&gt; 14

&lt;211&gt; 1563

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis sp.

&lt;400&gt; 14

atgtccgcc agatttcaat attccaagct cttgtctttc tattctaccg gtttatcctc 60  
 cggcgatata ggaactctaa accaaaatac caaaatggcc cttcttctct cctccaatcc 120  
 gacctatcac gccacacatt gatcttcaac gtagaaggag ctcttctcaa atccgactct 180  
 ctcttccctt acttcatgtt agtagcattt gaggcgggag gcgtaataag gtcatttctc 240  
 ctcttcattc tctatccatt gataagcttg atgagccatg agatgggtgt caaagtgatg 300  
 gtaatggtga gcttcttcgg gatcaaaaaa gaagggtttc gagcggggag agcgggtttg 360  
 cctaaatact ttctagaaga tgcgggactc gagatcttcg aagtgttgaa gagaggaggg 420  
 aagaaaatcg gaggagtgga tgatcttcct caagttatga tcgaagggtt cttgagagat 480  
 tacttgagga ttgacgttgt ggtcgggaga gaaatgaaag tcgttgaggg ttattatcta 540  
 ggtatcatgg aggataaaac caaacatgat cttgtctttg atgagttagt tcgtaaagag 600  
 agactaaaca ccggtcgtgt tattggcatc acttctctca atacatctct tcaccgatat 660  
 ctattctctc agttttgcc ggaattttat ttcgtgaaga aatcagacaa gcgaagctgg 720  
 caaaccttac cacgaagcca gtaccctaaa ccattgattt tccatgatgg ccgtctcgcg 780  
 atcaaaccac ccctaataaa cactttgggtc ttgttcattg ggggtccttt cgcagccgca 840  
 gccgcagcag ccagactctt cgtctctctt tgcacccctt actctttatc aatccccgatc 900  
 ctgccttttt ccggttgag actaacgctc actaacgact acgtttcatc tcaaaaacaa 960  
 aaaccaagtc aacgcaaagg ttgtctcttt gtatgtaacc ataggacttt attggaccct  
 1020  
 ctctatgttg cattcgcttt gagaaagaaa aacatcaaaa ctgtaacgta tagtttgagt  
 1080  
 agggatatctg agattttggc tccgatcaag acggtgagac tgaccctgta tcgggtgagc  
 1140

gacgggtcaag ccatggagaa attgttaacc gaaggagatc tcgttggttg tcctgaagga  
 1200  
 accacttgta gagaacctta cctgcttagg tttagccctt tggtcaccga ggtagtgat  
 1260  
 gtcacgttc ccgtggctgt gacgggtacac gtgaccttct tctacgggtac aacggcgagt  
 1320  
 ggtcttaagg cacttgaccc gcttttcttc ctcttggtac cttatcctac ctacaccatc  
 1380  
 caatttctcg accctgtctc cggtgccacg tgccaagatc ctgatggaaa gttgaagttt  
 1440  
 gaggtggcca acaatgttca gagtgatatt gggaaggcgc tggatttcga gtgcacaagt  
 1500  
 ctcactagaa aagacaagta tttgatcttg gccggtaata atggagtagt taagaaaaat  
 1560  
 taa  
 1563

<210> 15

<211> 520

<212> PRT

<213> Arabidopsis sp.

<400> 15

Met	Ser	Ala	Lys	Ile	Ser	Ile	Phe	Gln	Ala	Leu	Val	Phe	Leu	Phe	Tyr
1				5					10					15	
Arg	Phe	Ile	Leu	Arg	Arg	Tyr	Arg	Asn	Ser	Lys	Pro	Lys	Tyr	Gln	Asn
			20					25						30	
Gly	Pro	Ser	Ser	Leu	Leu	Gln	Ser	Asp	Leu	Ser	Arg	His	Thr	Leu	Ile
		35					40					45			
Phe	Asn	Val	Glu	Gly	Ala	Leu	Leu	Lys	Ser	Asp	Ser	Leu	Phe	Pro	Tyr
	50					55					60				
Phe	Met	Leu	Val	Ala	Phe	Glu	Ala	Gly	Gly	Val	Ile	Arg	Ser	Phe	Leu
	65				70					75					80
Leu	Phe	Ile	Leu	Tyr	Pro	Leu	Ile	Ser	Leu	Met	Ser	His	Glu	Met	Gly
				85					90					95	
Val	Lys	Val	Met	Val	Met	Val	Ser	Phe	Phe	Gly	Ile	Lys	Lys	Glu	Gly
			100					105						110	
Phe	Arg	Ala	Gly	Arg	Ala	Val	Leu	Pro	Lys	Tyr	Phe	Leu	Glu	Asp	Val
		115					120					125			
Gly	Leu	Glu	Ile	Phe	Glu	Val	Leu	Lys	Arg	Gly	Gly	Lys	Lys	Ile	Gly
	130					135					140				
Val	Ser	Asp	Asp	Leu	Pro	Gln	Val	Met	Ile	Glu	Gly	Phe	Leu	Arg	Asp
	145				150					155					160
Tyr	Leu	Glu	Ile	Asp	Val	Val	Val	Gly	Arg	Glu	Met	Lys	Val	Val	Gly
				165					170					175	
Gly	Tyr	Tyr	Leu	Gly	Ile	Met	Glu	Asp	Lys	Thr	Lys	His	Asp	Leu	Val
			180					185					190		
Phe	Asp	Glu	Leu	Val	Arg	Lys	Glu	Arg	Leu	Asn	Thr	Gly	Arg	Val	Ile
		195					200					205			
Gly	Ile	Thr	Ser	Phe	Asn	Thr	Ser	Leu	His	Arg	Tyr	Leu	Phe	Ser	Gln
	210					215					220				
Phe	Cys	Gln	Glu	Ile	Tyr	Phe	Val	Lys	Lys	Ser	Asp	Lys	Arg	Ser	Trp
	225				230					235					240
Gln	Thr	Leu	Pro	Arg	Ser	Gln	Tyr	Pro	Lys	Pro	Leu	Ile	Phe	His	Asp
				245					250					255	



Gly Arg Leu Ala Ile Lys Pro Thr Leu Met Asn Thr Leu Val Leu Phe  
                   260                  265                  270  
 Met Trp Gly Pro Phe Ala Ala Ala Ala Ala Ala Arg Leu Phe Val  
                   275                  280                  285  
 Ser Leu Cys Ile Pro Tyr Ser Leu Ser Ile Pro Ile Leu Ala Phe Ser  
                   290                  295                  300  
 Gly Cys Arg Leu Thr Val Thr Asn Asp Tyr Val Ser Ser Gln Lys Gln  
                   305                  310                  315                  320  
 Lys Pro Ser Gln Arg Lys Gly Cys Leu Phe Val Cys Asn His Arg Thr  
                   325                  330                  335  
 Leu Leu Asp Pro Leu Tyr Val Ala Phe Ala Leu Arg Lys Lys Asn Ile  
                   340                  345                  350  
 Lys Thr Val Thr Tyr Ser Leu Ser Arg Val Ser Glu Ile Leu Ala Pro  
                   355                  360                  365  
 Ile Lys Thr Val Arg Leu Thr Arg Asp Arg Val Ser Asp Gly Gln Ala  
                   370                  375                  380  
 Met Glu Lys Leu Leu Thr Glu Gly Asp Leu Val Val Cys Pro Glu Gly  
                   385                  390                  395                  400  
 Thr Thr Cys Arg Glu Pro Tyr Leu Leu Arg Phe Ser Pro Leu Phe Thr  
                   405                  410                  415  
 Glu Val Ser Asp Val Ile Val Pro Val Ala Val Thr Val His Val Thr  
                   420                  425                  430  
 Phe Phe Tyr Gly Thr Thr Ala Ser Gly Leu Lys Ala Leu Asp Pro Leu  
                   435                  440                  445  
 Phe Phe Leu Leu Asp Pro Tyr Pro Thr Tyr Thr Ile Gln Phe Leu Asp  
                   450                  455                  460  
 Pro Val Ser Gly Ala Thr Cys Gln Asp Pro Asp Gly Lys Leu Lys Phe  
                   465                  470                  475                  480  
 Glu Val Ala Asn Asn Val Gln Ser Asp Ile Gly Lys Ala Leu Asp Phe  
                   485                  490                  495  
 Glu Cys Thr Ser Leu Thr Arg Lys Asp Lys Tyr Leu Ile Leu Ala Gly  
                   500                  505                  510  
 Asn Asn Gly Val Val Lys Lys Asn  
                   515                  520

&lt;210&gt; 16

&lt;211&gt; 1506

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis sp.

&lt;400&gt; 16

atgggagctc aggagaaacg ggcgcgtttc gacgagatat caaagtgcga tgttaaggac 60  
 cgggtccaacc ataccgtggc cgctgatcta gacggaacac tactaatctc tcgtagcgcc 120  
 ttcccttact atttctcgt agccctcgag gcaggagct tgctccgagc gttgatccta 180  
 cttgtgtccg taccattcgt ttatcttacg tacttgacca tctccgagac tttagccatc 240  
 aacgtatttg tcttcatcac gttcgcgggt ctcaagatcc gagacgttga gctagtggtc 300  
 cgttccgtcc tcccagggtt ctatgcggag gacgtgaggc ccgatacctg gcgtatcttc 360  
 aacacgttcg ggaaacggta cataataact gcgagccctc gaattatggt cgagccattc 420  
 gtgaaaacat tcctaggagt tgataaagtt cttggaacag agctagaggt ctccaaatcg 480  
 ggtcggggcaa ccgggttcac cagaaaacca ggtattctcg tcggtcagta caaacgtgac 540  
 gtcgttttga gagagtgttg tggcctagcg tctgatttac ctgatttggg gctcggcgat 600  
 agcaagacgg accacgactt catgtccatc tgcaaggaag gttacatggt gccacgtacg 660

aaatgcgaac cattaccaag aaacaaactc ttaagcccca taatattcca cgagggcaga 720  
 ttagtccaac gcccaacgcc gttagtgtgt ctgttaactt tctctgtggt tcccgtcggt 780  
 ttcgtcctct ctatcatccg cgtctacacg aatattccgt taccggaacg tatcgcccgt 840  
 tacaactaca agcttactgg catcaagcta gtcgtcaacg gccaccctcc tccgcccga 900  
 aaacctggcc agccaggcca tcttttggtc tgcaaccacc gcaccgttct cgatcctgtg 960  
 gtcacagctg tcgcactcgg ccggaaaatc agctgcgtca cttacagcat cagcaagttc  
 1020  
 tctgagctaa tctcaccaat caaagccggt gcgttgactc gtcaacgtga gaaagacgca  
 1080  
 gcgaacatca agcgtctttt ggaggaaggc gatctcgtga tatgtcccga gggaaccacg  
 1140  
 tgccgtgagc ctttcttctt cgggtttagt gctcttttcg ctgagctcac ggaccggatc  
 1200  
 gttcccgtgg cgatcaacac aaagcagagc atgttcaatg gtaccaccac acgtggatac  
 1260  
 aagcttcttg atccttactt tgcgttcatt aacccgaggc cgacgtatga gatcacgttc  
 1320  
 ctcaaacaga ttccagctga gctgacgtgt aaaggaggca aatctccgat agagggttgcg  
 1380  
 aattacatac agagggtttt gggaggaacc ttaggttttg agtgacacaa tttcacaaga  
 1440  
 aaggataagt acgcaatgct tgctggtact gacggtaggg ttccgggtgaa gaaggagaag  
 1500  
 acgtga  
 1506

&lt;210&gt; 17

&lt;211&gt; 500

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis sp.

&lt;400&gt; 17

Met	Gly	Ala	Gln	Glu	Lys	Arg	Arg	Arg	Phe	Glu	Gln	Ile	Ser	Lys	Cys
1				5					10					15	
Asp	Val	Lys	Asp	Arg	Ser	Asn	His	Thr	Val	Ala	Ala	Asp	Leu	Asp	Gly
			20					25					30		
Thr	Leu	Leu	Ile	Ser	Arg	Ser	Ala	Phe	Pro	Tyr	Tyr	Phe	Leu	Val	Ala
		35					40					45			
Leu	Glu	Ala	Gly	Ser	Leu	Leu	Arg	Ala	Leu	Ile	Leu	Leu	Val	Ser	Val
	50						55				60				
Pro	Phe	Val	Tyr	Leu	Thr	Tyr	Leu	Thr	Ile	Ser	Glu	Thr	Leu	Ala	Ile
65					70					75					80
Asn	Val	Phe	Val	Phe	Ile	Thr	Phe	Ala	Gly	Leu	Lys	Ile	Arg	Asp	Val
			85						90					95	
Glu	Leu	Val	Val	Arg	Ser	Val	Leu	Pro	Arg	Phe	Tyr	Ala	Glu	Asp	Val
			100					105					110		
Arg	Pro	Asp	Thr	Trp	Arg	Ile	Phe	Asn	Thr	Phe	Gly	Lys	Arg	Tyr	Ile
		115					120					125			
Ile	Thr	Ala	Ser	Pro	Arg	Ile	Met	Val	Glu	Pro	Phe	Val	Lys	Thr	Phe
	130					135					140				
Leu	Gly	Val	Asp	Lys	Val	Leu	Gly	Thr	Glu	Leu	Glu	Val	Ser	Lys	Ser
145					150					155					160
Gly	Arg	Ala	Thr	Gly	Phe	Thr	Arg	Lys	Pro	Gly	Ile	Leu	Val	Gly	Gln
				165					170					175	
Tyr	Lys	Arg	Asp	Val	Val	Leu	Arg	Glu	Phe	Gly	Gly	Leu	Ala	Ser	Asp
			180					185					190		
Leu	Pro	Asp	Leu	Gly	Leu	Gly	Asp	Ser	Lys	Thr	Asp	His	Asp	Phe	Met
		195					200					205			

Ser Ile Cys Lys Glu Gly Tyr Met Val Pro Arg Thr Lys Cys Glu Pro  
 210 215 220  
 Leu Pro Arg Asn Lys Leu Leu Ser Pro Ile Ile Phe His Glu Gly Arg  
 225 230 235 240  
 Leu Val Gln Arg Pro Thr Pro Leu Val Ala Leu Leu Thr Phe Leu Trp  
 245 250 255  
 Leu Pro Val Gly Phe Val Leu Ser Ile Ile Arg Val Tyr Thr Asn Ile  
 260 265 270  
 Pro Leu Pro Glu Arg Ile Ala Arg Tyr Asn Tyr Lys Leu Thr Gly Ile  
 275 280 285  
 Lys Leu Val Val Asn Gly His Pro Pro Pro Pro Pro Lys Pro Gly Gln  
 290 295 300  
 Pro Gly His Leu Leu Val Cys Asn His Arg Thr Val Leu Asp Pro Val  
 305 310 315 320  
 Val Thr Ala Val Ala Leu Gly Arg Lys Ile Ser Cys Val Thr Tyr Ser  
 325 330 335  
 Ile Ser Lys Phe Ser Glu Leu Ile Ser Pro Ile Lys Ala Val Ala Leu  
 340 345 350  
 Thr Arg Gln Arg Glu Lys Asp Ala Ala Asn Ile Lys Arg Leu Leu Glu  
 355 360 365  
 Glu Gly Asp Leu Val Ile Cys Pro Glu Gly Thr Thr Cys Arg Glu Pro  
 370 375 380  
 Phe Leu Leu Arg Phe Ser Ala Leu Phe Ala Glu Leu Thr Asp Arg Ile  
 385 390 395 400  
 Val Pro Val Ala Ile Asn Thr Lys Gln Ser Met Phe Asn Gly Thr Thr  
 405 410 415  
 Thr Arg Gly Tyr Lys Leu Leu Asp Pro Tyr Phe Ala Phe Met Asn Pro  
 420 425 430  
 Arg Pro Thr Tyr Glu Ile Thr Phe Leu Lys Gln Ile Pro Ala Glu Leu  
 435 440 445  
 Thr Cys Lys Gly Gly Lys Ser Pro Ile Glu Val Ala Asn Tyr Ile Gln  
 450 455 460  
 Arg Val Leu Gly Gly Thr Leu Gly Phe Glu Cys Thr Asn Phe Thr Arg  
 465 470 475 480  
 Lys Asp Lys Tyr Ala Met Leu Ala Gly Thr Asp Gly Arg Val Pro Val  
 485 490 495  
 Lys Lys Glu Lys  
 500

&lt;210&gt; 18

&lt;211&gt; 1620

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis sp.

&lt;400&gt; 18

atggcgggac ctgatctgtc ttctcctttg atccaccatc aatcctccga tcaacctgaa 60  
 gttgttatct ctatcgccga cgacgacgac gacgagtcag gactcaatct tcttccagcc 120  
 gttgttgacc ctcgtgtttc acgaggtttt gagtttgacc atcttaatcc ttatggcttt 180  
 ctcagcgagt cagagcctcc ggttctcggt ccgacgacgg tggatccatt ccggaacaat 240  
 acacctggag ttagcggatt gtacgaagcg attaagctcg tgatttgtct tccgattgct 300

ctgattagac ttgttctctt tgctgctagc ttagctgttg gttacttggc tacaaaattg 360  
 gcacttgctg gctggaaaga taaagagaac cctatgcctc tttggagatg cagaatcatg 420  
 tggattactc ggatctgtac cagatgtatc ctcttctctt ttggctatca gtggataaga 480  
 aggaaaggga aacctgctcg gagagagatt gctccgattg ttgtatcaaa tcatgtttct 540  
 tatattgaac caatcttcta ctcttatgaa ttatcaccga ccattgttgc atcggagtca 600  
 catgattcac ttccatttgt tggaactatt atcagggcaa tgcaggtgat atatgtgaat 660  
 agattctcac agacatcaag gaagaatgct gtgcatgaaa taaagagaaa agcttctctg 720  
 gatagatttc ctctgtctgt gttattcccc gaaggaacca cgactaatgg gaaagtctct 780  
 atttcttccc aactcgggtgc tttcatccct ggttacccta ttcaacctgt agtagtccgg 840  
 tatccccatg tacattttga tcaatcctgg ggaaatatct ctttgttgac gctcatgttt 900  
 agaatgttca ctcatgttca caatttcatt gaggttgaat atcttctctg aatctatccc 960  
 agtgaaaaagc aaaagcagaa tgctgtgcgt ctctcacaga agactagtca tgcaattgca  
 1020  
 acatctttga atgtcgtcca aacatcccat tcttttgagg acttgatgct actcaacaaa  
 1080  
 gcaactgagt taaagctgga gaaccctca aattacatgg ttgaaatggc aagagttgag  
 1140  
 tcgctattcc atgtaagcag cttagaggca acgcgatttt tggatacatt tgtttccatg  
 1200  
 attccggact cgagtggacg tgttaggcta catgactttc ttcgggggtct taaactgaaa  
 1260  
 ccttgccctc tttctaaaag gatatttgag ttcacgatg tggagaaggc cggatcaatc  
 1320  
 actttcaaac agttcttgtt tgcctcgggc cacgtgttga cacagccgct ttttaagcaa  
 1380  
 acatgcgagc tagccttttc ccattgcat gcagatggag atggctatat tacaattcaa  
 1440  
 gaactcggag aagctctcaa aaacacaatc ccaaacttga acaaggacga gattcgagga  
 1500  
 atgtaccatt tgctagacga cgaccaagat caaagaatca gccaaaatga cttgttgtcc  
 1560  
 tgcttaagaa gaaacctctt tctcatagcc atctttgcac ctgacttggc cccaacataa  
 1620

&lt;210&gt; 19

&lt;211&gt; 539

&lt;212&gt; PRT

&lt;213&gt; Arabidopsis sp.

&lt;400&gt; 19

Met Ala Asp Pro Asp Leu Ser Ser Pro Leu Ile His His Gln Ser Ser  
 1 5 10 15

Asp Gln Pro Glu Val Val Ile Ser Ile Ala Asp Asp Asp Asp Asp Glu  
 20 25 30

Ser Gly Leu Asn Leu Leu Pro Ala Val Val Asp Pro Arg Val Ser Arg  
 35 40 45

Gly Phe Glu Phe Asp His Leu Asn Pro Tyr Gly Phe Leu Ser Glu Ser  
 50 55 60

Glu Pro Pro Val Leu Gly Pro Thr Thr Val Asp Pro Phe Arg Asn Asn  
 65 70 75 80

Thr Pro Gly Val Ser Gly Leu Tyr Glu Ala Ile Lys Leu Val Ile Cys  
 85 90 95

Leu Pro Ile Ala Leu Ile Arg Leu Val Leu Phe Ala Ala Ser Leu Ala  
 100 105 110

Val Gly Tyr Leu Ala Thr Lys Leu Ala Leu Ala Gly Trp Lys Asp Lys  
 115 120 125

Glu Asn Pro Met Pro Leu Trp Arg Cys Arg Ile Met Trp Ile Thr Arg  
 130 135 140

Ile Cys Thr Arg Cys Ile Leu Phe Ser Phe Gly Tyr Gln Trp Ile Arg  
 145 150 155 160

Arg Lys Gly Lys Pro Ala Arg Arg Glu Ile Ala Pro Ile Val Val Ser  
 165 170 175  
 Asn His Val Ser Tyr Ile Glu Pro Ile Phe Tyr Phe Tyr Glu Leu Ser  
 180 185 190  
 Pro Thr Ile Val Ala Ser Glu Ser His Asp Ser Leu Pro Phe Val Gly  
 195 200 205  
 Thr Ile Ile Arg Ala Met Gln Val Ile Tyr Val Asn Arg Phe Ser Gln  
 210 215 220  
 Thr Ser Arg Lys Asn Ala Val His Glu Ile Lys Arg Lys Ala Ser Cys  
 225 230 235 240  
 Asp Arg Phe Pro Arg Leu Leu Leu Phe Pro Glu Gly Thr Thr Thr Asn  
 245 250 255  
 Gly Lys Val Leu Ile Ser Phe Gln Leu Gly Ala Phe Ile Pro Gly Tyr  
 260 265 270  
 Pro Ile Gln Pro Val Val Val Arg Tyr Pro His Val His Phe Asp Gln  
 275 280 285  
 Ser Trp Gly Asn Ile Ser Leu Leu Thr Leu Met Phe Arg Met Phe Thr  
 290 295 300  
 Gln Phe His Asn Phe Met Glu Val Glu Tyr Leu Pro Val Ile Tyr Pro  
 305 310 315 320  
 Ser Glu Lys Gln Lys Gln Asn Ala Val Arg Leu Ser Gln Lys Thr Ser  
 325 330 335  
 His Ala Ile Ala Thr Ser Leu Asn Val Val Gln Thr Ser His Ser Phe  
 340 345 350  
 Ala Asp Leu Met Leu Leu Asn Lys Ala Thr Glu Leu Lys Leu Glu Asn  
 355 360 365  
 Pro Ser Asn Tyr Met Val Glu Met Ala Arg Val Glu Ser Leu Phe His  
 370 375 380  
 Val Ser Ser Leu Glu Ala Thr Arg Phe Leu Asp Thr Phe Val Ser Met  
 385 390 395 400  
 Ile Pro Asp Ser Ser Gly Arg Val Arg Leu His Asp Phe Leu Arg Gly  
 405 410 415  
 Leu Lys Leu Lys Pro Cys Pro Leu Ser Lys Arg Ile Phe Glu Phe Ile  
 420 425 430  
 Asp Val Glu Lys Val Gly Ser Ile Thr Phe Lys Gln Phe Leu Phe Ala  
 435 440 445  
 Ser Gly His Val Leu Thr Gln Pro Leu Phe Lys Gln Thr Cys Glu Leu  
 450 455 460  
 Ala Phe Ser His Cys Asp Ala Asp Gly Asp Gly Tyr Ile Thr Ile Gln  
 465 470 475 480  
 Glu Leu Gly Glu Ala Leu Lys Asn Thr Ile Pro Asn Leu Asn Lys Asp  
 485 490 495  
 Glu Ile Arg Gly Met Tyr His Leu Leu Asp Asp Asp Gln Asp Gln Arg  
 500 505 510  
 Ile Ser Gln Asn Asp Leu Leu Ser Cys Leu Arg Arg Asn Pro Leu Leu  
 515 520 525  
 Ile Ala Ile Phe Ala Pro Asp Leu Ala Pro Thr

530

535

<210> 20  
<211> 1128  
<212> DNA  
<213> Arabidopsis sp.

<400> 20  
atggaaaaaa agagtgtacc aaattctgat aagttgtctc tgattagagt gttaagaggt 60  
ataaatatgtc tgatggtgtt agtttcaaca gcttttatga tgttgatatt ctggggggttc 120  
ttatcagctg tagtggttag gcttttcagc attcgctata gccgtaaatg tgtttccttc 180  
ttctttggct cgtggctcgc cttgtggcct ttctctttg agaagattaa caaaaccaa 240  
gttatcttct ctggtgataa ggttccttgc gaggatcgag tattgctcat tgcaaaccac 300  
cgaacagaag ttgattggat gtacttctgg gatcttgcac tgcgtaaagg ccagattggg 360  
aatatcaaatt atgtgcttaa gagtagtttg atgaaattac ctctctttg ttgggcgttt 420  
cacctctttg agtttattcc tgttgagagg agatgggaag tcgatgaagc aaacttgaga 480  
cagatcttct cgagttttaa ggatccccga gagcctttat ggcttgctct ttccccgag 540  
ggcacagatt acacagaggc taaatgccaa aggagtaaga aatttgctgc tgaaaatggc 600  
cttccgatac tgaacaacgt gctgcttccc aggacaaaag gtttcgtctc ctgcttgcaa 660  
gaactgagtt gctcacttga cgcagtttat gatgtgacca tcggttataa aaccgcgtgc 720  
ccatcttctc tagacaacgt ttatggaatt gagccatcag aagttcacat ccacatccgt 780  
cgtatcaacc tgacccaaat cccaaatcaa gaaaaggaca tcaatgcttg gttaatgaac 840  
acattccagc tcaaagacca gctgctcaat gacttttact ccaatgggtca ttccctaac 900  
gaaggaacag agaaagagtt caacacaaag aagtacctca taaactgttt ggcagtgatt 960  
gccttcacca ccatctgtac acatctcacc ttcttctcat caatgatttg gttcaggatt  
1020  
tatgtctctt tggcctgtgt ctacttgacc tctgctacgc atttcaatct tcgttctggt  
1080  
ccactgtgtg agactgcaaa aaattccctc aaattagtaa acaaataa  
1128

<210> 21  
<211> 375  
<212> PRT  
<213> Arabidopsis sp.

<400> 21  
Met Glu Lys Lys Ser Val Pro Asn Ser Asp Lys Leu Ser Leu Ile Arg  
1 5 10 15  
Val Leu Arg Gly Ile Ile Cys Leu Met Val Leu Val Ser Thr Ala Phe  
20 25 30  
Met Met Leu Ile Phe Trp Gly Phe Leu Ser Ala Val Val Leu Arg Leu  
35 40 45  
Phe Ser Ile Arg Tyr Ser Arg Lys Cys Val Ser Phe Phe Phe Gly Ser  
50 55 60  
Trp Leu Ala Leu Trp Pro Phe Leu Phe Glu Lys Ile Asn Lys Thr Lys  
65 70 75 80  
Val Ile Phe Ser Gly Asp Lys Val Pro Cys Glu Asp Arg Val Leu Leu  
85 90 95  
Ile Ala Asn His Arg Thr Glu Val Asp Trp Met Tyr Phe Trp Asp Leu  
100 105 110  
Ala Leu Arg Lys Gly Gln Ile Gly Asn Ile Lys Tyr Val Leu Lys Ser  
115 120 125  
Ser Leu Met Lys Leu Pro Leu Phe Gly Trp Ala Phe His Leu Phe Glu  
130 135 140  
Phe Ile Pro Val Glu Arg Arg Trp Glu Val Asp Glu Ala Asn Leu Arg  
145 150 155 160  
Gln Ile Val Ser Ser Phe Lys Asp Pro Arg Asp Ala Leu Trp Leu Ala  
165 170 175

Leu Phe Pro Glu Gly Thr Asp Tyr Thr Glu Ala Lys Cys Gln Arg Ser  
 180 185 190  
 Lys Lys Phe Ala Ala Glu Asn Gly Leu Pro Ile Leu Asn Asn Val Leu  
 195 200 205  
 Leu Pro Arg Thr Lys Gly Phe Val Ser Cys Leu Gln Glu Leu Ser Cys  
 210 215 220  
 Ser Leu Asp Ala Val Tyr Asp Val Thr Ile Gly Tyr Lys Thr Arg Cys  
 225 230 235 240  
 Pro Ser Phe Leu Asp Asn Val Tyr Gly Ile Glu Pro Ser Glu Val His  
 245 250 255  
 Ile His Ile Arg Arg Ile Asn Leu Thr Gln Ile Pro Asn Gln Glu Lys  
 260 265 270  
 Asp Ile Asn Ala Trp Leu Met Asn Thr Phe Gln Leu Lys Asp Gln Leu  
 275 280 285  
 Leu Asn Asp Phe Tyr Ser Asn Gly His Phe Pro Asn Glu Gly Thr Glu  
 290 295 300  
 Lys Glu Phe Asn Thr Lys Lys Tyr Leu Ile Asn Cys Leu Ala Val Ile  
 305 310 315 320  
 Ala Phe Thr Thr Ile Cys Thr His Leu Thr Phe Phe Ser Ser Met Ile  
 325 330 335  
 Trp Phe Arg Ile Tyr Val Ser Leu Ala Cys Val Tyr Leu Thr Ser Ala  
 340 345 350  
 Thr His Phe Asn Leu Arg Ser Val Pro Leu Val Glu Thr Ala Lys Asn  
 355 360 365  
 Ser Leu Lys Leu Val Asn Lys  
 370 375

&lt;210&gt; 22

&lt;211&gt; 1170

&lt;212&gt; DNA

&lt;213&gt; Arabidopsis sp.

&lt;400&gt; 22

atgggtgattg ctgcagctgt catcgtgcct ttgggccttc tcttcttcat atctgggtctc 60  
 gctgtcaatc tctttcaggc agtttgctat gtactcattc gaccactgtc taagaacaca 120  
 tacagaaaaa ttaaccgggt gggtgcagaa accttggtgt tggagcttgt atggatagtt 180  
 gactgggtggg ctggagttaa gatccaagtg tttgctgata atgagacctt caatcgaatg 240  
 ggcaaagaac atgctcttgt cgtttgtaat caccgaagtg atattgattg gcttgtggga 300  
 tggattcttg ctcagcggtc aggttgcttg ggaagcgcat tagctgtaat gaagaagtct 360  
 tccaaattcc ttccagtcac aggctgggtca atgtggttct cggagtatct ctttctggaa 420  
 agaaattggg ccaaggatga aagcactcta aagtcaggtc ttcagcgctt gagcgacttc 480  
 cctcgacctt tctggttagc cctttttgtg gagggaaactc gctttacaga agccaaactt 540  
 aaagccgcac aagagtatgc agcctcctct gaattgccta tccctcgaaa tgtgttgatt 600  
 cctcgcacca aagggttctgt gtcagctggt agtaatatgc gttcatttgt ccagcaatt 660  
 tatgatatga cagtgactat tccaaaaacc tctccaccac ccacgatgct aagactattc 720  
 aaaggacaac ctctagttgt gcatgttcac atcaagtgtc actcgatgaa agactttacc 780  
 gaatcagatg acgcaattgc acagtgggtg agagatcagt ttgtggctaa ggatgctctg 840  
 ttagacaaac acatagctgc agacactttc cccggccaac aagaacagaa cattggccgt 900  
 cccataaagt cccttgcggt ggttctatca tgggcatgcg tactaactct tggagcaata 960  
 aagttcctac actgggcaca actcttttct tcatggaaag gtatcacgat atcggcgctt  
 1020  
 ggtctaggta tcatcactct ctgtatgcag atcctgatac gctcgtctca gtcagagcgt  
 1080  
 tcgaccccag ccaaagtcgt ccagccaag ccaaagaca atcaccaccc agaatcatcc  
 1140

tcccaaacag aaacggagaa ggagaagtaa  
1170

<210> 23

<211> 389

<212> PRT

<213> Arabidopsis sp.

<400> 23

Met	Val	Ile	Ala	Ala	Ala	Val	Ile	Val	Pro	Leu	Gly	Leu	Leu	Phe	Phe
1				5					10					15	
Ile	Ser	Gly	Leu	Ala	Val	Asn	Leu	Phe	Gln	Ala	Val	Cys	Tyr	Val	Leu
			20					25					30		
Ile	Arg	Pro	Leu	Ser	Lys	Asn	Thr	Tyr	Arg	Lys	Ile	Asn	Arg	Val	Val
		35					40					45			
Ala	Glu	Thr	Leu	Trp	Leu	Glu	Leu	Val	Trp	Ile	Val	Asp	Trp	Trp	Ala
	50					55					60				
Gly	Val	Lys	Ile	Gln	Val	Phe	Ala	Asp	Asn	Glu	Thr	Phe	Asn	Arg	Met
65					70					75					80
Gly	Lys	Glu	His	Ala	Leu	Val	Val	Cys	Asn	His	Arg	Ser	Asp	Ile	Asp
				85					90					95	
Trp	Leu	Val	Gly	Trp	Ile	Leu	Ala	Gln	Arg	Ser	Gly	Cys	Leu	Gly	Ser
			100					105					110		
Ala	Leu	Ala	Val	Met	Lys	Lys	Ser	Ser	Lys	Phe	Leu	Pro	Val	Ile	Gly
		115					120					125			
Trp	Ser	Met	Trp	Phe	Ser	Glu	Tyr	Leu	Phe	Leu	Glu	Arg	Asn	Trp	Ala
	130					135					140				
Lys	Asp	Glu	Ser	Thr	Leu	Lys	Ser	Gly	Leu	Gln	Arg	Leu	Ser	Asp	Phe
145					150					155					160
Pro	Arg	Pro	Phe	Trp	Leu	Ala	Leu	Phe	Val	Glu	Gly	Thr	Arg	Phe	Thr
			165					170						175	
Glu	Ala	Lys	Leu	Lys	Ala	Ala	Gln	Glu	Tyr	Ala	Ala	Ser	Ser	Glu	Leu
			180					185						190	
Pro	Ile	Pro	Arg	Asn	Val	Leu	Ile	Pro	Arg	Thr	Lys	Gly	Phe	Val	Ser
		195					200					205			
Ala	Val	Ser	Asn	Met	Arg	Ser	Phe	Val	Pro	Ala	Ile	Tyr	Asp	Met	Thr
	210					215					220				
Val	Thr	Ile	Pro	Lys	Thr	Ser	Pro	Pro	Pro	Thr	Met	Leu	Arg	Leu	Phe
225					230					235					240
Lys	Gly	Gln	Pro	Ser	Val	Val	His	Val	His	Ile	Lys	Cys	His	Ser	Met
				245					250					255	
Lys	Asp	Leu	Pro	Glu	Ser	Asp	Asp	Ala	Ile	Ala	Gln	Trp	Cys	Arg	Asp
			260					265					270		
Gln	Phe	Val	Ala	Lys	Asp	Ala	Leu	Leu	Asp	Lys	His	Ile	Ala	Ala	Asp
		275					280					285			
Thr	Phe	Pro	Gly	Gln	Gln	Glu	Gln	Asn	Ile	Gly	Arg	Pro	Ile	Lys	Ser
	290					295					300				
Leu	Ala	Val	Val	Leu	Ser	Trp	Ala	Cys	Val	Leu	Thr	Leu	Gly	Ala	Ile
305					310					315					320
Lys	Phe	Leu	His	Trp	Ala	Gln	Leu	Phe	Ser	Ser	Trp	Lys	Gly	Ile	Thr



325 330 335  
 Ile Ser Ala Leu Gly Leu Gly Ile Ile Thr Leu Cys Met Gln Ile Leu  
 340 345 350  
 Ile Arg Ser Ser Gln Ser Glu Arg Ser Thr Pro Ala Lys Val Val Pro  
 355 360 365  
 Ala Lys Pro Lys Asp Asn His His Pro Glu Ser Ser Ser Gln Thr Glu  
 370 375 380  
 Thr Glu Lys Glu Lys  
 385

<210> 24  
 <211> 269  
 <212> DNA  
 <213> Glycine max

<400> 24  
 gaccactga acgctctcat caccttcacg tggtccccc tgggttcat cctctccatc 60  
 ataagggctt acttcaacct cctctccca gaacncattg tccgtacac ctacgagatg 120  
 ctcgccatca acctcgatc cgcgggccac cgcctctctc cgccttccc cgccacccc 180  
 ggcaacctct acgtctgcaa ccaccgcacc gctctcgacc ccacgtcat cgccattgcc 240  
 ctcgcccgca aggtctcctg cgtcaccta 269

<210> 25  
 <211> 242  
 <212> DNA  
 <213> Glycine max

<400> 25  
 tgatcttcca cgacggccgt ttcgtgcaga ggccagaccc actgaacgct ctcatcacct 60  
 tcacgtggct ccccttcggc ttcacacctt ccatacataa ggtctacttc aaccttcttc 120  
 tcccagaacg cattgtccgc tacacctacg agatgctcgg catcaacctc gtcattccgcg 180  
 gccaccgccc tcttccgctt tcccccgga cccccggcaa cctctacgtc tgcaaccacc 240  
 gc 242

<210> 26  
 <211> 272  
 <212> DNA  
 <213> Glycine max

<400> 26  
 gtttgttcaa aggccaactc ctctagcagc cctcttgacc ttcctatggt tgccaattgg 60  
 catcatactc tccatnctta agggctctacc ttaacatccc ttgacctgaa agaattgctt 120  
 ggtataacta taagctatta ggaatcagag ttattgtgaa ggggtaccct ccaccacccc 180  
 caaagaaggg tcaaagtggg gtcctatttg tttgtaacca ccgcacagtt ttagaccctg 240  
 tggttactgc agttgcactt ggaagaaaaa tt 272

<210> 27  
 <211> 218  
 <212> DNA  
 <213> Glycine max

<400> 27  
 atagcacagg agggttacat ggtgcctccg agcaaatacag caaaggcagt cccacaggag 60  
 cgtctgaaga gcagaatgat cttccacgac gggcggttctg tgcagaggcc agacccaatg 120  
 aatgccctca tcaccttcac atggctccct ttgggttctg tcctctccat cataagggtc 180  
 tacttcaacc tccctctccc agaacgcacg gtcgcgta 218

<210> 28  
 <211> 270  
 <212> DNA  
 <213> Glycine max

<400> 28  
 gtgcctgttg ctgtgaactg caagcagaac atgttctttg gaaccaccgt tcgtggcgctc 60  
 aagttctggg acccttaact tacttcttac atgaacccta ggcctgtgta cgagggttacc 120

ttaccttgat acctttgccg aggagatgtc ggttaaggct ggggggaagt cgtccattga 180  
ggtggccaac cacgtggcag aaggtgctgg gggatgtgtt agggtttgag tgcaccgggt 240  
tgactaggaa ggataagtat atgttggttg 270

<210> 29  
<211> 252  
<212> DNA  
<213> Glycine max

<400> 29  
catgagggtta ggtttgctca aaggccaact cctctagctg ccctcttgac cttcctatgg 60  
ctgccaattg gcatacact ctccatctta agggctacc ttaacatccc ttgacctgaa 120  
agaattgttg gtacaactac aagctcttag gaatcagagt tattgtgaag ggtaccctc 180  
caccgccccc aaagaagggt caaagtgggt tctatttgtt tgtaaccacc gcacagtatt 240  
agaccctgtt gt 252

<210> 30  
<211> 272  
<212> DNA  
<213> Glycine max

<400> 30  
ctgggactgc cttaaactgat gcattggatct tatcaagaaa ggagcctctg tttttttctt 60  
tccagagggga acacgcagta aagatggaag actaggcaca ttcaagaagg gtgctttcag 120  
tggtgctgca aagacaaatg caccagtagt accaattacc cttattggaa ctgggtcaaa 180  
catgcctgca ggaaaggagg gaatagtga cataggttct gtgaaagtgg ttatacataa 240  
acctattgtt ggaaaggatc ctgacatgtt at 272

<210> 31  
<211> 239  
<212> DNA  
<213> Glycine max

<400> 31  
cggaatcaa ggtcatcaga cttcaagggt gtttcagctg ttgtcactga cagaattcga 60  
gaagctcatc agaatgagtc tgctccatta atgatgttat ttccagaagg tacaaccaca 120  
aatggagagt tcctccttcc attcaagact ggtggttttt tggcaaaggc accgggtactt 180  
cctgtgatat tacgatatca ttaccagaga tttagccctg cctgggattc catatctgg 239

<210> 32  
<211> 242  
<212> DNA  
<213> Glycine max

<400> 32  
gaacggcaac ggcaacagcg ttcgcgatga cgcctcctctg ctgaagccgg agcctccggt 60  
cttcgcgcga cagcatcgcc gatatggaga agaagttcgc cgcttacgtc cgccgctacg 120  
tgtacggcac catgggacgc ggcgagttgc ctcccaagga gaagctcttg ctcggtttcg 180  
cgttggtcac tcttctcccc attcgagtcg ttctcgccgt caccatattg ctcttttatt 240  
ac 242

<210> 33  
<211> 248  
<212> DNA  
<213> Glycine max

<400> 33  
ttcttcttct ctactctctt aaaaccctaa ctctatacat ggaagggaata nctcaaattct 60  
natgactaat taattaatcc atcgatcaag catggagtcc gaactcaaag acctcaattc 120  
gaagccgccc aacggcaacg gcaacagcgt tcgcgatgac cgtcctctgc tgaagccgga 180  
gcctccggtc tccgcccaga gcacgcccga tatggagaag aagttcgccg cttacgtccg 240  
ccgcgacg 242

<210> 34  
<211> 217  
<212> DNA  
<213> Glycine max

<400> 34  
aaaaccctaa ttctatacat ggaagggaata tctcaaattct aatgactaat taattaatcc 60

```

atcgatcaag catggagtcc gaactcaaag acctcaattc gaagccgccg aacggcaacg 120
gcaacagcgt tcgcatgac cgtcctctgc tgaagccgga gcctccggtc tccgccgaca 180
gcatcgccga tatggagaag aagttcgccg cttacgt 217

```

<210> 35  
 <211> 257  
 <212> DNA  
 <213> Glycine max

```

<400> 35
atctctgtct ctgcatttcc ctccctaaaa ccctaattct acatttggaa aggaaatctc 60
aaatctaata actaattaat caatcaatcg tattaataat ccatcgatca agtatggagt 120
ccgaactcaa agacctcaat tcgaagccac ccaactgcaa cggaacgcc aacagcggtt 180
gagacgaccg tcctctgctg aagcgggagc ctccggcctc ctccgacagc atcgccgaga 240
tggagaagaa gttcgcc 257

```

<210> 36  
 <211> 284  
 <212> DNA  
 <213> Glycine max

```

<400> 36
cccgacaaaa acagggtttt gtggccaatc atacttccat gattgatttc attatcttag 60
aacagatgac tgcatttgct gttattatgc agaagcatcc tggatgggtt ggattattgc 120
agagcaccat tntggagagt gtaggggtgta tctggttcaa ccgtacagag gcaaaggatc 180
gagaagttgt ggcaaggaaa ttgagggatc atgtcctggg agctaacaac aacctctctc 240
ttatatctcc tgaaggaact tgtgtaata atcactactc gtca 284

```

<210> 37  
 <211> 246  
 <212> DNA  
 <213> Glycine max

```

<400> 37
ggagatccgc ataagcaaata caatcctcct gttccttcct tatctctgtc tctgcatttc 60
cctccctaaa accctaattc tacatttgga aaggaantct caaatctaata gataattaat 120
caatcaatcg tattaataat ccatcgatca agtatggagt ccgaactcaa agacctcaat 180
tcgaagccac ccaactgcaa cggaacgcc aacagcggtt gcgacgaccg tcctctgctg 240
aagccg 246

```

<210> 38  
 <211> 278  
 <212> DNA  
 <213> Glycine max

```

<400> 38
gttttcttatt gccacgttgt ggaagcgtaa cgaagatgaa tggcattggg aaactcaaata 60
cgtcgagtgc tgaattggac cttcacattg aagattacct accttctgga tccagtgttc 120
aacaagaacg gcatggcaag ctccgactgt gtgatttgct agacatttct cctagtctat 180
ctgaggcagc acgtgccatt gtagatgata cattcacaag gtgcttcaag caaatcctcc 240
agaaccttgg aactggaatg tttatttgtt tcctttgt 278

```

<210> 39  
 <211> 312  
 <212> DNA  
 <213> Glycine max

```

<400> 39
ttaactttgg cacattctcc ttttgttcat caatgtgtgt tgtaaattgt ncatttcctt 60
cagaggtctt tggtaganat gatgtgcagt ttctgtgggt catcttggac tngngntgtt 120
aagnatcatg gaccaggcc tagcaggaga ccaaagcagg ttttgttagc caaccatact 180
tcatgattga tntcattatn tnagaacaga tgactgcttt tgcngttatn atgcagaagc 240
atcctggatg ggttggttaag cntacagnat gtcaacngtg tatnaaatat gntacacnnn 300
acttgcgtct tc 312

```

<210> 40  
 <211> 255  
 <212> DNA  
 <213> Glycine max

&lt;400&gt; 40

```

ggattattgn ngcanatgca gtcattctgtt ctaagataat ganatcnatc atggaagtat 60
gattggnacac anaaacctgt yttttgggtg gatactagggt cttggcccat ggtacttgac 120
naccctcagtc catgatgcaa canaganact gnacatcatc tccaccaaac ccctctgana 180
ganacgagaa ttgagcaatt tagagtacct tggtttgatg caagtcagta tattcaagtt 240
tctattcatc aaagg                                     255

```

&lt;210&gt; 41

&lt;211&gt; 291

&lt;212&gt; DNA

&lt;213&gt; Glycine max

&lt;400&gt; 41

```

caacctccca tgcaatcgct caccctctcc gtcacctgaa tctgttttct attccctccg 60
tcgcgtaaca aggatgaatg gcattgggaa actcaaactcg tcgagttctg aattggacct 120
tcacattgaa gattacctgc cttctggatc cagtgttcaa caagaacggc atggcaagct 180
ccgcctgtgt gatttgctag acatttctcc tagtctatct gaggcagcac gtgccattgt 240
agatgataca ttcacaagggt gtttcaagtc aaatcctcca gaaccttgga a          291

```

&lt;210&gt; 42

&lt;211&gt; 284

&lt;212&gt; DNA

&lt;213&gt; Glycine max

&lt;400&gt; 42

```

ctgcaacctta ccatgcaatt cctcacctga atccgttttc tattgccacg ttgtggaagc 60
gtaacgaaga tgaatggcat tgggaaactc aaatcgctga gttctgaatt ggaccttcac 120
attgaagatt acctaccttc tggatccagt gttcaacaag aacggcatgg caagctccga 180
ctgtgtgatt tgctagacat ttctcctagt ctatctgagg cagcacgtgc catgtagatg 240
atacatcaca aggtgctcaa gtcaaacttc cagaaccttg gaat          284

```

&lt;210&gt; 43

&lt;211&gt; 268

&lt;212&gt; DNA

&lt;213&gt; Glycine max

&lt;400&gt; 43

```

ctgaagtatt ctcgtcctag cccaaagcat agagaaaggn agcaacagaa ctttgctgag 60
tcagtgtctgc ggcgatggga ggaaaagtga tgtgtacctt tatgtggtgt tgttcttaat 120
tattcttagt aatgccattg cttcgacccc tttttttgct tttgttttgt cattgctaac 180
tattttattt taacactttt attaaagata tggcatatat ncacttcagt anacaaagtt 240
gtncacagtaa tttnttttcc aaaaaaaaa          268

```

&lt;210&gt; 44

&lt;211&gt; 241

&lt;212&gt; DNA

&lt;213&gt; Glycine max

&lt;400&gt; 44

```

gancaaaatt gccctccatc actttccttg ttagagttgg tttctgcnac ctaccatgca 60
attccctcac ctgaatccgt tttctattgc cactgtgtgg aagcgtaacg aagatgaatg 120
gcattgggaa actcaaactg tcgagttctg aattggacct tcacattgaa gattacctac 180
cttctggatc cagtgttcaa caagaacggc atggcaagct ccgactgtgt gatttgctag 240
a                                     241

```

&lt;210&gt; 45

&lt;211&gt; 247

&lt;212&gt; DNA

&lt;213&gt; Glycine max

&lt;400&gt; 45

```

gtaggatgtc tgagatcctt gcccacatca aaacggtgcg gttaactaga aaccgcgacg 60
aggatgcgaa aatgatgaaa aatttgctgg ggcaagggga cctgggtggt tgtcctgaag 120
ggaccacatg tagagaacct tatttattga ggttcagccc tctgttctca gagatgtgag 180
atgagattgt ccccggtggc agttgattcc cagttatatg ttccacggaa ccactgctgg 240
tgganta                                     247

```

&lt;210&gt; 46

&lt;211&gt; 271

&lt;212&gt; DNA

<213> Glycine max

<400> 46

```
tgcagggggg cttgttagag ccatagtttt ggttcttcta tacccttttg tttgtgtcgt 60
aggaaaagag atgggggttga agataatggt catggcatgc ttcttcggga tcaaagcatc 120
gagcttcaga gttggaaggt ccgttttgcc cnaattcttc tnggaggacg ttngtgcaga 180
aatgtttgag gcactcaaaa aaggaggga gacagtggga gttaccaatt taccacacgt 240
gatggtggaa agcttcttga gagagtattt g                                     271
```

<210> 47

<211> 242

<212> DNA

<213> Glycine max

<400> 47

```
ttcacagctg tcacgccgtn aacggaaaat ggcaacggcg agacgcagtt tcccgctat 60
caccgaatgc aacggaaacga cncggtgga ntctgtngnc gccgacctcg agggtagcgt 120
cctcatctcc cgtngctcgt tcccgtaatt catgctcgtc gccgtcgaag ccggcagcnt 180
cctccggggc ctcatgctnc tctctccct tccgttcgtc atnategcct acctcttcat 240
ct                                     242
```

<210> 48

<211> 244

<212> DNA

<213> Glycine max

<400> 48

```
acatattctt cagttagctc ccccaaccta tacacttcac caccacacca caaccctacc 60
ctctctctct gtcatgggtca ttggaggagc cttccctcgt ttcgacccaa tcaccaaatg 120
tagacccaag accgctccaa ccagaccatc gcctcggacc tcgatggcac cctccttgtc 180
tcccggaagt cttccacct ctacttctc gtcgccctcg aagccggcag cgtcttccga 240
gcct                                     244
```

<210> 49

<211> 230

<212> DNA

<213> Glycine max

<400> 49

```
caacattcca cctagctccc caatcacatc ttcaccacac cataaacctt cttaatttct 60
ctcttcattt tctcctctat tgtcataatc atggggacct tccctcgtt cgacccaatc 120
accaccaag accggtccaa ccagaccgtg gcctccgacc ttgacggcac cctcctcgtc 180
tcccggaagc cttccacct ctacctctc gttgccctcg aagccggcag                                     230
```

<210> 50

<211> 265

<212> DNA

<213> Glycine max

<400> 50

```
ctggtgaata atcctaagtt atggagtctg tgggtgtgtga gctagaaggc acgcttgtga 60
aggacaagga tgcgttctca tacttcatgt tgggtgcgtt tgaagcttca ggtttggttc 120
gtttcgctt gttgctaaca ctattgcccg tgattcgggt ccttgacatg gttggcatga 180
acgatgcac tctcaagcta ntnatcttcg tggctgtggc tgggtgtcca aagtccgaga 240
ttgaatcagt ggctagggca gtttt                                     265
```

<210> 51

<211> 252

<212> DNA

<213> Glycine max

<400> 51

```
ctggtgaata atcctaagtt atggagtctg tgggtgtgtga gctagaaggc acgcttgtga 60
aggacaagga tgcgttctca tacttcatgt tgggtgcgtt tgaagcttca ggtttggttc 120
gtttcgctt gttgctaaca ctattgcccg tgattcgggt ccttgacatg gttggcatga 180
acgatgcac tctcaagcta atgatcttcg tggctgtggc tgggttccaa agtccgagat 240
tgaatcagt gc                                     252
```

<210> 52

<211> 218

&lt;212&gt; DNA

&lt;213&gt; Glycine max

&lt;400&gt; 52

aactgcaact	acaacaacat	tcattcattc	acagctgtca	cgccgtgaac	ggaaaatggc	60
aacggcgaga	cgcagtttac	ccgcctatac	accgaatgca	acggaacgac	accgtgcgag	120
tctgtggcgc	ccgacctcga	cggtacgctc	ctcatntccc	gtagctcggt	cccgtacttc	180
atgctcgtcg	ccgtcgaagc	cggcagcctc	ctccgcgg			218

&lt;210&gt; 53

&lt;211&gt; 262

&lt;212&gt; DNA

&lt;213&gt; Glycine max

&lt;400&gt; 53

ggttaaggac	attgagatgg	tcgnntcctc	ggtgctgccc	aagttctaca	ccgaggacgt	60
gcnccccag	agctggagag	tcttcaatcc	ttcgggaagc	gttacattgt	cactgctagt	120
ctagggtgat	ggtggagcan	tttggttaaga	cgtttcttgg	ggctgataag	gtgcttgagg	180
ctgagcttga	ggccacgaaa	tcggggaggt	tcatgggttt	gttaaggagc	ctggtgtgct	240
tgttggggag	cacaagaaag	tg				262

&lt;210&gt; 54

&lt;211&gt; 212

&lt;212&gt; DNA

&lt;213&gt; Glycine max

&lt;400&gt; 54

gcaactacaa	caacattcat	tcattcacag	ctgtcacgcc	gtgaacggaa	aatggcaacg	60
gcgagacgca	gtttcccgcg	tatcacggaa	tgcaacggaa	cgacgccgtg	cgagtctgtg	120
gccgccgacc	tcgacggtac	gtcctctatc	tcccgtagnc	cgttcccgtg	cttcatgctc	180
gtngccgtcg	aagccggcag	cctcctccgc	gg			212

&lt;210&gt; 55

&lt;211&gt; 273

&lt;212&gt; DNA

&lt;213&gt; Glycine max

&lt;400&gt; 55

catggttttc	ttgagcttct	ttggcctcag	aaaggacaca	ttcagaacag	gatcagctgt	60
tctggcaaa	ttcttcttag	aagatgttgg	attggaaggc	tttgaggccg	taatatgttg	120
tgagagaaaa	gtggcatcta	gtaagttgcc	aagggtcatg	gttgaaaatt	tcctcaagga	180
ctatttaggg	gttgatgctg	ttatagcaag	agaattgaag	tccttttagt	gcttcttttt	240
gggagttttt	gagagtaaga	agccaattaa	aat			273

&lt;210&gt; 56

&lt;211&gt; 257

&lt;212&gt; DNA

&lt;213&gt; Glycine max

&lt;400&gt; 56

ctctcaaaaa	aggagggaag	acagtgggag	tcaccaatct	accccatgtg	atggtggaaa	60
gcttcttgag	agagtatttg	gacattgatt	tcgttgtggg	cagggagctg	aaagttttct	120
gtggatacta	cgtaggattg	atggatgaca	caaaaactat	gcatgccttg	gagctgggta	180
aagaaggaaa	aggatgctcc	gacatgatcg	gaatcacaag	gtttcgcaac	atacgcgacc	240
atgatgattt	tttctcc					257

&lt;210&gt; 57

&lt;211&gt; 240

&lt;212&gt; DNA

&lt;213&gt; Glycine max

&lt;400&gt; 57

gaactaagtg	tgaaccacta	ccaagaaaca	agcttttaag	tccaattatt	tttcatgagg	60
gtaggtttgc	tcaaaggcca	actcctctag	ctgnnctctt	gaccttccta	tggctgccaa	120
ttggcatcat	actctccatc	ttaagggctc	accttaacat	ccctttgcct	gaaagaattg	180
cttggtacaa	ctacaagctc	ttaggaatca	gagttattgt	gaagggtacc	cctccaccgc	240

&lt;210&gt; 58

&lt;211&gt; 254

&lt;212&gt; DNA

&lt;213&gt; Glycine max

&lt;400&gt; 58

```
cttggaataa ggggtcattag gaaggggtatc cctccacccc cagcnaagaa gggccaaagt 60
ggagtcctat ttgtatgcaa ccacaggaca gtttttagacc ctgtgggttac agctgttgca 120
ttaggaagga aaattagctg tgtcacatat agcataagca aattcactga aataatttca 180
ccaatcaaag ctgtggcact ctctagggag agggacaaag atgctgccaa catcaagang 240
ttgcttgagg aagg                                     254
```

&lt;210&gt; 59

&lt;211&gt; 267

&lt;212&gt; DNA

&lt;213&gt; Glycine max

&lt;400&gt; 59

```
gccaganaga cttgcttggt acaactacaa gcttcttgga ataaggggtca ttaggaaggg 60
tatccctcca cccccagcaa agaagggcca aagtggagtc ctatttgtat gcaaccacag 120
gacagtttta gacctgtggt ttacagctgt tgcattagga aggaaaatta gctgtgtcac 180
atatagcata agcaaattca ctgaaataat tcaccaatca aagctgtggc actctctagg 240
gagagggacc nagatgctgc cnacatc                                     267
```

&lt;210&gt; 60

&lt;211&gt; 261

&lt;212&gt; DNA

&lt;213&gt; Glycine max

&lt;400&gt; 60

```
gtaaccacag ggtctaaaac tgtgcggtgg ttactgcagt tgcacttgnc nagaaaaatt 60
tgcttatgct atatgtgaca cagctaattc actgnaataa tttcaccaat taaagctgtg 120
gcactctcaa ggganngaga gaaagatgct gccaatatcc ngagactact tgaggaaggg 180
gacttggtga tttgcctga aggcacaact tgtagagagc cttcctcttg aggttcagtg 240
cactatttgc tgaactcact g                                     261
```

&lt;210&gt; 61

&lt;211&gt; 258

&lt;212&gt; DNA

&lt;213&gt; Glycine max

&lt;400&gt; 61

```
caaggagctc acatgcagtg gagggaaatc agctattgaa gttgcaaact acattcaaag 60
ggttcttgca gggacttttg gatttgagtg cacaaatttg actaggaaga gcaaatatgc 120
catgcttgca ggcacagatg ggacagttcc atctaaggag aaggcttgan aaggggagaga 180
aattaagttc tcccttttga ttattctgta ttggtgccca atgtgtttcc aaaacactta 240
gaattatgat agaaataa                                     258
```

&lt;210&gt; 62

&lt;211&gt; 258

&lt;212&gt; DNA

&lt;213&gt; Glycine max

&lt;400&gt; 62

```
attggcataa tcctctccat cctaagggtc tatctcaaca tccctctgcc agaaagactt 60
gcttgntaca actacaagct tcttggaata agggtcatta ggaaggggtat ccctccaccc 120
ccagcaaaga agggccaaag tggagcctat ttgtatgcaa ccacaggaca gtttttagacc 180
ctgtgggttac agctgttgca ttaggaagga aaattagctg tgtcacatat agcataagca 240
aattcactga aataattt                                     258
```

&lt;210&gt; 63

&lt;211&gt; 239

&lt;212&gt; DNA

&lt;213&gt; Glycine max

&lt;400&gt; 63

```
cacttcacca ccacaccaca accctaccct ctctctctgt catgggtcatt ggaggagcct 60
tccctcgttt cgacccaatc accaaatgta gcacccaaga ccgctccaac cagaccatcg 120
cctcggacct cgatggcacc ctccctgtct ccgggagtgc cttcccctac tacttccctcg 180
tcgccctcga agccggcagc gtcttcctag ccctccttct cttaaccttc gtccccttc 239
```

&lt;210&gt; 64

&lt;211&gt; 531

&lt;212&gt; DNA

&lt;213&gt; Glycine max

&lt;400&gt; 64

```

cggagaaccg gtctaacc aa accgtggcct cggacttgga cggcaccctc ctggtgtccc 60
ccagcgcatt tccttactac atgctgggtcg ccacogaagc cggcagcttc ctccgtggcc 120
ttgtcctcct tgccctcgtc ccttttcgtgt attcacgtac atattcctct ccgagaccgc 180
ggccatcaag tccctgatct tcatcgctt cgcgggcctg aaggtcaggg acgttgagat 240
ggtcgcgtgc tcggtgctgc ccaagttcta cgccgacata ttcttcagtt agtccccca 300
acctatacac ttcaccacca caccacaacc ctaccctctc tctctgtcat ggtcattgga 360
ggagccttcc ctcggttctga cccaatcacc aaatgtagca cccaagaccg ctccaaccag 420
accatcgctt cggacctcga tggcaccctc cttgtctccc ggagtgcctt ccctactac 480
ttcctcgtcg ccctcgaagc cggcagcgtc ttccgagccc tccttctctt a 531

```

&lt;210&gt; 65

&lt;211&gt; 256

&lt;212&gt; DNA

&lt;213&gt; Glycine max

&lt;400&gt; 65

```

acatattctt cagttagctc ccccaacctt tacacttcac caccacacca caaccctacc 60
ctctctctct gtcattgtca ttggaggagc cttccctcgt ttcgacccaa tcaccaaagt 120
tagcacccaa gaccgctcca accagaccat cgcctcggac ctcgatggca ccctccttgt 180
ctcccgaggt gccttccctt actacttctt cgtcgccttc gaagccggca gcgtcttccg 240
agccctcctt ctctta

```

&lt;210&gt; 66

&lt;211&gt; 260

&lt;212&gt; DNA

&lt;213&gt; Glycine max

&lt;400&gt; 66

```

ccatccaaca tattcttcag ttagctcccc caacctatac acttcaccac cacaccacaa 60
ccctaccctc tctctctgtc atgggtcattg gaggagcctt ccctcggttc gacccaatca 120
ccaaatgtag cacccaagac cgctccaacc agactatcgc ctcgacctc gatggcacc 180
tccttgtctc ccggagtgc ttccccctact acttctcgt cgcctcga gccggcagcg 240
tcttcgagc cctccttctc

```

&lt;210&gt; 67

&lt;211&gt; 248

&lt;212&gt; DNA

&lt;213&gt; Glycine max

&lt;400&gt; 67

```

caccaaccaa acctcactct ccttttctcc cctgaccctc tccctgcat ggtcatggga 60
gcctttggcc acttcgaacc ggtctccaaa tgcagcaccg agaaccgggt taaccaaac 120
gtggcctcgg acttgacgg caccctcctg gtgtcccca gcgcatttcc ttactacatg 180
ctgggcgcca tcgaagccgg cagcttctct cgtggccttg tcctccttgc ctccgtccct 240
ttcgtgta

```

&lt;210&gt; 68

&lt;211&gt; 283

&lt;212&gt; DNA

&lt;213&gt; Glycine max

&lt;400&gt; 68

```

ttcttcccca ccatcacacc aancaaacct cactctnctt ggccatgggt atgnnngcct 60
ttccggccact tcgaaccggt ttccaaatgc agcaccgaaa accgggttaa ccaaacctgt 120
gcctcggact tggacggcac cctcctgggt tcccctagcg cctttcctta ctacatgctc 180
gtcgccatcg aagccggcag cttcctcctg ggccttgtcc tccttggatc cgtcccttcc 240
gtgtacttca cgtacatatt cttctccgag accgcggcca tca 283

```

&lt;210&gt; 69

&lt;211&gt; 258

&lt;212&gt; DNA

&lt;213&gt; Glycine max

&lt;400&gt; 69

```

ctctttcttcc ccaccatcnn accaaccaaa cctcactctc cctgaccatg gtcattgggag 60
cctttcgcca cttcgaaccg gtttccaaat gcagcaccga aaaccgggtt aaccaaaccg 120

```



tggcctcgga cttggacggc accctcctgg tgccccctag cgcctttcct tactacatgc 180  
tcgtcgccat cgaagcggc agcttcctcc gtggccttgt cctccttga tccgtccctt 240  
tcgtgtactt cacgtaca 258

<210> 70  
<211> 256  
<212> DNA  
<213> Glycine max

<400> 70  
tgcaactaca acaacattca ttcattcaca gctgtcacgc cgtgaacgga aaatggcaac 60  
ggcgagacgc agtttcccg cttaccgga atgcaacgga acgacaccgt gcgagtctgt 120  
ggcgcggac ctcgacggta cgtcctcat ctcccgtagc tcgttcccg acttcatgct 180  
cgtcgccgtc gaagcggca gntcctccg cggcctcatc ctctcctng ccantccgtt 240  
cgtcatcanc gcctac 256

<210> 71  
<211> 259  
<212> DNA  
<213> Glycine max

<400> 71  
cttccccacc atcacaccan ggcnacctc antctccctt tctccacnga cctctcctt 60  
gccatngtca tgggancctt tggccacttc gaaccggtct ccaaagcag caccgagaac 120  
cggntaacc aaaccgtggc ctccgacttg gacggcacc tcttggtgtc ccncagcgca 180  
tttcttact acatgctggc ngccatcgaa gccggcagct tctcctgtg ccttgctctc 240  
cttgctccg tcccttctg 259

<210> 72  
<211> 249  
<212> DNA  
<213> Glycine max

<400> 72  
ccaacatatt cttcagttag ctcccccaac ctatacactt caccaccaca ccacaacctt 60  
accctctctc tctgtcatgg tcattggagg agccttccct cgtttcgacc caatcaccaa 120  
atgtagcacc caagaccgct ccaaccagac catcgctcgc gacctcgatg gaccctnct 180  
tgtctcccg agtgccttc cctactactt cctcgtcgcc ctogaagcgc gcagcgtctt 240  
ncgagccct 249

<210> 73  
<211> 257  
<212> DNA  
<213> Glycine max

<400> 73  
caaccctctt cttccccacc atcacaccaa ncaaacctca ctctccctt ctccccctgac 60  
cctctccctg ccatggctat gggagcctt ggccacttcg aaccggtctc caaatgcagc 120  
accgagaacc ggtctaacca aaccgtggcc tcggacttgg acggcaccct cctggtgtcc 180  
cccagcgcat ntccttacta catgctggtc gccatcgaag ccggcagctt cctccgtggc 240  
cttgctctcc ttgctg 257

<210> 74  
<211> 255  
<212> DNA  
<213> Glycine max

<400> 74  
gccgaagacg tgcacccgga gagttggaga gtgttcaact ctttcgggaa gcgttacatt 60  
gtcacggcta gtctagggg gatgggtggag ccgtttgtta aggcgtttct cggggctgac 120  
aagggtgctt ggactgaact tgaggccacc aaatcgggga cgttactgg gtttgtaag 180  
aagcctggtg tgcttgttgg ggagcataag aaagtggctc tgggtgaagga gtttcagggt 240  
aattacctga cttgg 255

<210> 75  
<211> 244  
<212> DNA  
<213> Glycine max

<400> 75

```

caacaacatt cattcattca cagctgtcac gccgtgaacg gaaaatggca acggcgagac 60
gcagtttccc gcctatcacc gaatgcaacg gaacgacacc gtgcgagtc gtggccgccc 120
acctcgacgg tacgtctctc atccccgta gctcgttccc gtacttcctg ctgcgtgccc 180
tcgaagccgg cagcctctc cgcggcctca tgcnttcctg gggtttanttt gagnaccctt 240
gagg

```

<210> 76  
 <211> 240  
 <212> DNA  
 <213> Glycine max

```

<400> 76
gctggctacc ctcttcttcc ccaccatcac accaatcaaa cctcactcta ccctggccat 60
ggtcatggga gcctttncgc cacttcgaac cgggtttccaa atgcagcacc gaanaccggg 120
ttaccanac cgtggcctcg gncttggacg gcaccctcct ggtgtcccct agcgcctttc 180
cttactacat gctcgtcgcc atcgaagccg gcagcttcct ccgtggcttg tctccttgg 240

```

<210> 77  
 <211> 263  
 <212> DNA  
 <213> Glycine max

```

<400> 77
gtttctcggg gctgacaagg tgcttgggac tgaacttgag gccaccaa atcggggacgtt 60
cactgggttt gttaagaagc ctggtgtgct tggttggggag cataagaaag tggctctggt 120
gaaggagttt cagggttaatt tacctgactt ggggtctaggt gatagtaaaa gtgattatga 180
cttcatgtca atttgcaagg aaggggtacat ggtgccaaga actaagtgtg aaccactacc 240
aagaacaag cttttaagtc caa

```

<210> 78  
 <211> 258  
 <212> DNA  
 <213> Glycine max

```

<400> 78
ggccacgaaa tcggggaggt tcaactgggt tgtaaggag cctgggtgtgc ttgttgggga 60
gcacaagaaa gtggctgttg tgaaggaggt tcagggtaat ttacctgact tgggactagg 120
agatagtaaa agtgattatg acttcatgtc aatttgcaag gaagggtaca tgggtccaag 180
gactaagtgt gaaccactac caagaaaca acttttaagt ccaattattt ntcattgagg 240
taggtttgtt caaaggcc

```

<210> 79  
 <211> 260  
 <212> DNA  
 <213> Glycine max

```

<400> 79
ctcttcttcc ccaccatcac accaancaaa cctcactctc cctttctccc ctgaccctct 60
ccctgccatg gtcatgggag cctttggcca cttcgaaccg gtctccaaat gcagaccgga 120
gaaccggctc aaccaaaccg tggcctcgga cttggacggc accctcctgg tgtccccag 180
cgcatttcct tactacatgc tggtcgccat cgaagccggc agcttccctc gtgggccttg 240
tctccttgc ctccgtccct

```

<210> 80  
 <211> 257  
 <212> DNA  
 <213> Glycine max

```

<400> 80
gggaacaaca acaaatggca ngaaccttat ctctttccaa cttggtgcat ttatccctgg 60
ataccaatc cagcctgtaa ttgtacgcta tctcatgtg cactttgacc aatcctgggg 120
tcatgtntct ttgggaaagc ttatgttcag aatgttcact caatttcaca acttttttga 180
ggtagaatat ctctctgtca tttatccccct ggatgataag gaaactgctg tancctntcg 240
ggagaggact agccggg

```

<210> 81  
 <211> 272  
 <212> DNA  
 <213> Glycine max

<400> 81  
 catacctttt gttggcacca ttattagagc aatgcaggtc atatatgtta acagattctt 60  
 accatcatca aggaagcagg ctgttaggga aataaaggaa ctgaataaca gagaagggcc 120  
 tcttgtgata aatttcctcg agtactatta tttcccagg gaacaacaac taatggcagg 180  
 aaccttatct ccttccaact tgggtcattt atccctggat acccaatcca gcctgtaatt 240  
 atacgctatc ctcatgtaca ctttgaccaa tc 272

<210> 82  
 <211> 245  
 <212> DNA  
 <213> Glycine max

<400> 82  
 gggcatttca catactagag ttcattccag tgaaaagaaa gtgggaggct gatgaatcaa 60  
 tcattgcgcca tatgttttct acattcaagg atccacaaga tcctctctgg ctgctgcttt 120  
 tcccagaagg cactgatttc actgagcaaa agtgccctcg gagtcaaaaa tatgtctgtg 180  
 aacataagtt accggttctg aaaaatgttt tacttccaag gacaaagggg cttctgtgcc 240  
 gcttg 245

<210> 83  
 <211> 268  
 <212> DNA  
 <213> Glycine max

<400> 83  
 cagtgtcctt cctttctgga caatgttttt ggtgttgacc cttcagaagt gcacctgcat 60  
 gtgctggcgta ttccgggtgga ggagattcca gcttctgaaa ccaaagctgc ttcttggtta 120  
 atcgacacat tccagatcaa ggaccaattg ctttcggatt tcaagattca aggccatttc 180  
 cctaaccaac taaatgaaaa tgaaatttct agatttaaga gcctactctc ttttatggtg 240  
 atagtttctt ttactgccat gtttattt 265

<210> 84  
 <211> 265  
 <212> DNA  
 <213> Glycine max

<400> 84  
 gaaagagact gggcaaaaaga tgaaacatca ctgaagtcag gttttaggca tctagagcac 60  
 atgccattcc ctttctggtt ggcccttttt gttgaaggaa ctcgtttcac gcagacaaaag 120  
 cttttacaag ctcaagagtt tgctgcttca aaagggtgc ctatacctag aaatgttttg 180  
 attcctcgta ctaagggttt tgtcacagca gnacaaagcc ttcggccatt tcgttccagc 240  
 catttatgat tgcacatatg cagtt 265

<210> 85  
 <211> 265  
 <212> DNA  
 <213> Glycine max

<400> 85  
 gaaagagact gggcaaaaaga tgaaacatca ctgaagtcag gttttaggca tctagagcac 60  
 atgccattcc ctttctggtt ggcccttttt gttgaaggaa ctcgtttcac gcagacaaaag 120  
 cttttacaag ctcaagagtt tgctgcttca aaagggtgc ctatacctag aaatgttttg 180  
 attcctcgta ctaagggttt tgtcacagca gnacaaagcc ttcggccatt tcgttccagc 240  
 catttatgat tgcacatatg cagtt 265

<210> 86  
 <211> 301  
 <212> DNA  
 <213> Zea mays

<400> 86  
 ctctgctgca agggcaccct gccgcccgcg cccaagaagg gccaccggg cgtcctcttc 60  
 gtctgcaacc accgcaccgt gctcgaccct gtcgaggttg ccgtggcgct gcgcccgaag 120  
 gtcagctgcg tcacctacag catctccaag ttctccgagc tcattctgce catcaaggcc 180  
 gtccgctgtg cgcgggagge gacaaggacg ccgagaacat ccgcccctg ctggaggagg 240  
 gcgacctggt catctgcccc gagggnaaca actgccgcga gcccttctct ctgcgttcag 300  
 g 301

<210> 87  
 <211> 309

&lt;212&gt; DNA

&lt;213&gt; Zea mays

&lt;400&gt; 87

```

cgctcatgcg gtgtacatca acctgccgct gcccgagcgc atcgtctact acacctacaa 60
gctcatgggc atcaggctcg tcgtcaaggg caccgccg ccgcccgcga agaagggcca 120
cccgggcgct ctcttcgtct gcaaccaccg caccgtgctc gaccccgctc aggtggccgt 180
ggcgctgcgc cgcaaggtca gctgcgtcac ctacagcatc tccaagttct ccgagctcat 240
ctcgcccatc aaggccgctc cgctgtcggg gaggcgacaa ggacgccgag aacatccgcc 300
gcctgctgg 309

```

&lt;210&gt; 88

&lt;211&gt; 304

&lt;212&gt; DNA

&lt;213&gt; Zea mays

&lt;400&gt; 88

```

tggctgtgca ggaggcctac ctggtgacgt caaggaagta cagcccggtg cccaggaacc 60
agctgctgag ccgctgatt cgtgcacgac ggccgcctcg tgcagcgccc gacgccgctc 120
gtcgcgctcg tcaccttct ctggatgccg ttcggttcg cgctggcgct catgcgcgtg 180
tacatcaacc tgccgctgcc cgagcgcatc gtctactaca cctacaagct catgggcac 240
aggctcgctg tcaagggcac ccgcccgcg cgcgccaaga agggccacc gggcgctcctc 300
ttcg 304

```

&lt;210&gt; 89

&lt;211&gt; 312

&lt;212&gt; DNA

&lt;213&gt; Zea mays

&lt;400&gt; 89

```

ggttcatcca cttgtgttgc tattngaccg gtaccgtagg agagcacagc actancatcg 60
caaagatttn gggctacggt gacaatctcc atgttctaca atcttnaggt cgaaggaatg 120
gagaatctgc ctccaaatag ctgtcctggt gtctatgttg ctaaccatca gagcttcttg 180
gatatttata cccttctaac tctagggagg tgcttcaaat ttataagcaa gaccagcatc 240
tttatgttcc ctattatagg gtgggcaatg tatctcttgg gtgtgattcc tctgcggcgt 300
atggacagca gg 312

```

&lt;210&gt; 90

&lt;211&gt; 264

&lt;212&gt; DNA

&lt;213&gt; Zea mays

&lt;400&gt; 90

```

ggtgctgtat ctgaaagaat ccatcgctgt catcaacaga aaaatgcacc aatgatgcta 60
ctcttcccct gagggcacaa ctacaaatgg ggattatctc cttccattca aaacaggtgc 120
ttttcttgca aaggcaccag ttcaaccagt cattttgaga tatccttaca aaagatttaa 180
tgcagcatgg gattccatgt caggggcacg tcatgtattt ctgctgctct gtcaatttgt 240
aaattaccta gaggtggtcc gctt 264

```

&lt;210&gt; 91

&lt;211&gt; 212

&lt;212&gt; DNA

&lt;213&gt; Zea mays

&lt;400&gt; 91

```

aaatgtcttg gatgcatttt tggtcagcgg gagtcgaaaa caccagattt caaagggtgtt 60
tcaggtgctg tatttgaaag aatccatcgt gctcatcaac agaaaaatgc accaatgatg 120
ctactcttcc ctgagggcac aactacaaat ggggattatc tccttccatt caaaacaggt 180
gcttttcttg caaaggcacc agttcaacca gt 212

```

&lt;210&gt; 92

&lt;211&gt; 267

&lt;212&gt; DNA

&lt;213&gt; Zea mays

&lt;400&gt; 92

```

gtctaaagaa atngaaaggc gtggggnaat tgtgtctaata catgtncttt atgtggatat 60
tctttatcan atgtcagcct ctttccctag tttgttgct aagagatcag tggntagatt 120
gcctctagtt ggtctcataa gcaaatgtct tggatgcatt tttgttcagc gggagtnnaa 180
aatncanatt tcaaagggtg ttaagggtgt gnatctgaaa gaatccatcg tgctcatcaa 240

```

cagaaaaatg caccaatgat gctactc

267

&lt;210&gt; 93

&lt;211&gt; 152

&lt;212&gt; DNA

&lt;213&gt; Zea mays

&lt;400&gt; 93

ctacaaatgg	ggattacctt	cttccattta	agactggagc	ctttnttgca	ggtgcaccag	60
tgcagccagt	cattttgaaa	tacccttaca	ggagatttag	tccagcatgg	gattcaatgg	120
atggagcacg	tcattgtgta	ttgctgctct	gt			152

&lt;210&gt; 94

&lt;211&gt; 274

&lt;212&gt; DNA

&lt;213&gt; Zea mays

&lt;400&gt; 94

aaaatataaa	ttaatatggt	cttaatccca	ccatataaat	aacgttctct	ttctgcaggg	60
caatttagtt	ctttctaata	ttgggctggc	agagaagcgc	gtgtaccatg	cagcactgac	120
tggtagtagt	ctacctggcg	ctagacatga	gaaagatgat	tgaaagacgt	tgcgctcgctt	180
tttctgtaac	agacagccga	ggaacactta	aaaatgtaac	tgtgtgcgtg	tttttatacc	240
tgtaatgtgg	cagtttattt	gtttgaggag	gctg			274

&lt;210&gt; 95

&lt;211&gt; 295

&lt;212&gt; DNA

&lt;213&gt; Zea mays

&lt;400&gt; 95

aatagctatc	aagtacaata	aaatatttgt	tgatgccttt	tggaacagta	agaagcaatc	60
ttttacaatg	cacttggtcc	ggctgatgac	atcatgggct	gttgtgtgtg	atgtttggta	120
cttacctcct	caatatctga	gggagggaga	gacggcaatt	gcatttgctg	agagagtaag	180
ggacatgata	gctgctagag	ctggactaaa	gaaggttcct	tgggatggct	atctgaaaca	240
caaccgtcct	agtcccaaac	acactgaaga	gaacaacgca	tattgccgat	ctgtc	295

&lt;210&gt; 96

&lt;211&gt; 273

&lt;212&gt; DNA

&lt;213&gt; Zea mays

&lt;400&gt; 96

gngccatctc	accggcggn	ggcctgcggc	cggcaaccgg	aggcgatggc	gagctngtct	60
gtggtggcgg	acatggagca	ntaccgcccc	aacctggagg	actacctccc	gcccgactcg	120
ctcccgcagg	aggcgcccag	gaatctccat	ctgcgcgatc	tgcttgacat	ctcgccggtg	180
ctaaccgagg	cagcggggtg	catagtcgat	gattcattca	cccgttgctt	taagtccaat	240
tctccagaac	catggaatgg	aacatatatt	tgt			273

&lt;210&gt; 97

&lt;211&gt; 127

&lt;212&gt; DNA

&lt;213&gt; Zea mays

&lt;400&gt; 97

ctcaatatct	ganggagga	gagactgcaa	ttgcgtttgc	tgagagagta	agggacatga	60
tagcagctag	agctggtctt	aagaaggctc	cgtgggatgg	ctatctgaag	cacaaccgcc	120
ctagtcc						127

&lt;210&gt; 98

&lt;211&gt; 286

&lt;212&gt; DNA

&lt;213&gt; Zea mays

&lt;400&gt; 98

gaaccgtacg	cgcctcatta	cgcccatcca	cgtgctcgcc	tctccccatc	gcataatttt	60
nctcggcggc	gtcgccatct	ccanccggng	cnggcctgcn	gccggcaacc	ggaggcgatg	120
gcgagctcgt	ctgtggcggc	ggacatggag	ctggaccgcc	ccaacctgga	ggactacntc	180
ccgcccagant	cgctcccga	ggaggcgacc	aggaatctcc	atctgngcga	tctgcttgan	240
atctcgccgg	tgctaaccga	ggcagcgggg	gccatagtcg	atgatt		286

<210> 99  
<211> 308  
<212> DNA  
<213> Zea mays

<400> 99  
cgccatctca tcggcggcgg gcgtgcggcc ggcggcngag gcgaggngcg attggcgagc 60  
tcgtctgtgg cgccggacat ggagctggac cgcccanacc tggaggacta nctcccggcc 120  
gactcgnncc cgcagaggcg ccccggaatc tccanctgcy cgatctgctg gacatcncgc 180  
cgggtgtcac cgaggcagcg ggtgccattg tcgatgactc cttcacacgg ngctttaagt 240  
caaattctcc agagccatgg aattggaaca tatatctgtt ccccttatgt gctttgggtg 300  
ataataag 308

<210> 100  
<211> 282  
<212> DNA  
<213> Zea mays

<400> 100  
cagaaactag angttagtca cagcatggca ttaaattgtc atagtaaaca acancncact 60  
gagcaactat gcaatttaac gccatgctgt gactaacttc tagtttctgg cattaataa 120  
ctgtttggct actaggaaga ccgaggtaga gaagcaaata taagaatacc ctccaacgca 180  
canccaaatg acagagtaaa tgaaggtagg gtacaccttc ttgaacatga ccgtatactg 240  
gttgtaaaca caagttcctc tgggaaaatc agagaggggt tt 282

<210> 101  
<211> 282  
<212> DNA  
<213> Zea mays

<400> 101  
ggcgcggtcg gccgtggcgc tggtcctgcc gtacagtact cgacgccgat cctggcngcg 60  
acnggcattg cgtggcggtc caaagggtn ggcncngnc ttgcnngcc gtgctccggc 120  
ggcgctgnc agctgttcgt gtgcaacnac cggacgctga tcgaccngt gtacgtgtcc 180  
gtagcgtgga ccgggaaatg cgcnctgtgt nctacagnct gangcggntn tcggagctca 240  
tctcccccat ngncggaang tgcacctgan accgggaacg gg 282

<210> 102  
<211> 290  
<212> DNA  
<213> Zea mays

<400> 102  
ggacgcggca ccatgcgcgc cgagctggcc agtggcgacg tggccgtgtg ccccgagggc 60  
accacgtgcc gggagccctt cctgctccgc ttctccaagc tcttcgcgga gctcagcgac 120  
aggatcgtgc ccgtggcgat gaactaccgc gtggggctct tccaccgcac gacggcgcg 180  
gggtggaaag ccatggaccc catcttcttc ttcatgaacn gcggcccgtg tacgaggtga 240  
cgttcctgaa ccantccccg caaagcgacg tgcgcggcgg ggaagagccc 290

<210> 103  
<211> 279  
<212> DNA  
<213> Zea mays

<400> 103  
acgaggtgac gttcctgaac cagctccccg cagaggcgac gtgcgcggcg gggaagagcc 60  
ccgttgatgt agccaactac gttcagcgga tactcgctgc cacgctcggg ttcgagtga 120  
ccaccctcac aaggaaggac aaatacacgg tgctcgccgg caacgacggc gtcctgaacg 180  
ccaagccggc ggcgggcccg aagccggctt ggcagagccg cgtgaaggaa gtcctcgggt 240  
tctgctccac taacaattac accttgccca gatctggac 279

<210> 104  
<211> 315  
<212> DNA  
<213> Zea mays

<400> 104  
gccccgagcg atcgtctact acacctacaa gctcatgggc atcaggctcg tcgtcaaggg 60  
caccgcggcg ccgcccgcga agaagggcca cccgggcgtc ctcttcgtct gcaaccaccg 120  
caccgtgctc gaccccgctg aggtggccgt ggcgtgcgc cgcaangtca gtcgctcac 180

tacagcatct ccaagttctc cgagctcatc tcgcccata aggccgtagc agnaaagcag 240  
 gtcgcaaatg gagcagnagc gagtcgatgg aagngaattg gcgactgggc atctgcncga 300  
 aggnacactg cggag 315

<210> 105

<211> 314

<212> DNA

<213> Zea mays

<400> 105

cgagacaccg agcacgtact accagcaaga tgggtggcgtc tcccagattc aagcccatcg 60  
 aggagtgtctg ctccggagggg cggtcggagc agacgggtggc cgccgacctg gacggcacgc 120  
 tgctcatctc caggagcgcg tccccctact acctcctcgt ggctctcgag gccggcagcg 180  
 tcttcgcgcg cgcgctgtg ctctgttccg tgcggttcgt ctacgtcacc tacgccttct 240  
 tctccgagtc gctggccatc agcacgctgg tgtacatctc cgtggcgggg ctcaagggtgc 300  
 gcanatcgag atgg 314

<210> 106

<211> 291

<212> DNA

<213> Zea mays

<400> 106

ctctgggtct gggggccgaga caccgagcac gtactaccag caagatgggtg gcgtctccca 60  
 gattcaagcc catcgaggag tgctgtctcg aggggcggtc ggagcagacg gtggccgccc 120  
 acctggacgg cacgctgtct atntccagga gcgcttccc ctactacctc ctctgtggctc 180  
 tcgaggccgg cagcgtctc cgcgcgcgcg tgctgtctct gtccgtgccc ttcgtctacg 240  
 tcacctacgc cttcttctcc gagtcgctgg ccatcagcac gctgggtgtac a 291

<210> 107

<211> 300

<212> DNA

<213> Zea mays

<400> 107

gcacgcagca gtacgacgtc tctcctctgg gtctggggcc gagacaccga gcacgtacta 60  
 ccagcaagat ggtggcgtct cccagattca agcccatcga ggagtgtgctc tcggagggggc 120  
 ggtcggagca gacgggtggcc gccgacctgg acggcacgct gctcatctcc aggagcgcgt 180  
 tcccctacta cctcctcgtg gctctcgagg ccggcagcgt cctccgcgcc gcgctgctgc 240  
 tctgttccgt gcggttcgtc tacgtcacct acgccttctt ctccgagtcg ctggccatca 300

<210> 108

<211> 284

<212> DNA

<213> Zea mays

<400> 108

gnggccgaga caccgagcac gtactaccag cagatgggtg gcgtctccca gattcangcc 60  
 antcaggagg tgctgtctcg aggggcggtc ggagcagacg gtggccgccc acctggacgg 120  
 cacgctgtct atctccagga gcgcttccc ctacnacctc ctctgtggctc tcgaggccgg 180  
 cagcgtctct cgcgcgcgcg tgctgtctct gtccgtgccc ttcgtctacg tcactacgcc 240  
 ttcttctccg agtcgctggc catcaanacg ctgggtgtaca tctc 284

<210> 109

<211> 280

<212> DNA

<213> Zea mays

<400> 109

ctcctctggg tctggggccg agacaccgag cacgtactac cagcaagatg gtggcgtctc 60  
 ccagattcaa gcccatcgag gagtgctgct cggagggggc gtcggagcag acgggtggccg 120  
 ccgacctgga cggcacgctg ctcatctcca ggagcgcgtt ccnctactac ctctcgtgg 180  
 ctctcgaggg cggcagcgtc ctccgcgcgg cgtgtgtgt cctgtccgtn ccgttcgtct 240  
 acgtcaccta cgcnttnttc tccgagtcgc tggccatcag 280

<210> 110

<211> 287

<212> DNA

<213> Zea mays

<400> 110  
 cgtctctcct ctgggtctgg ggccgagaca ccgagcacgt actaccagca agatgggtggc 60  
 gtctcccaga ttcaagccca tcgaggagtg ctgctcggag gggcggtcgg agcagacggc 120  
 ggccgccgac ctggacggca gctgctcatc tccaggagcg cgttccccta ctacctctc 180  
 gtggctctcg aggcggcgag cgtctccgc gccgcgctgc tgcctctgtc cgtgcggttc 240  
 gtctacgtca ctacggcttc ttctccgagt cgtgggcat cagcacg 287

<210> 111  
 <211> 286  
 <212> DNA  
 <213> Zea mays

<400> 111  
 cgcacagtta cgacgtctct cctctgggtc tggggccgag acaccgagca cgtactacca 60  
 gcaagatggc ggcgtctccc agattcaagc ccatcgagga gtgctgctcg gagggcggt 120  
 cggagcagac ggtggccgcc gacctggacg gcacgctgct catctccagg agcgcgttcc 180  
 cctactactc ctgctgctct cgaggccggc aggtcctccg cgccgcgctg tgcctctgtc 240  
 gtgcgttcgt ctagtacta cgtttttctc gancgtggca ataana 286

<210> 112  
 <211> 323  
 <212> DNA  
 <213> Zea mays

<400> 112  
 gttattccct gaaggtacca caacaaatgg gagattcctg atttcgttcc aacatgggtc 60  
 attcatacct ggctaccctg ttcaacctgt tgttgtccgt tatccacatg tgcactttga 120  
 tcaatcatgg gggnatatat cgttattaaa gctcatgttt aagatgttca cccaatttca 180  
 taatttcatg gaggtagagt accttcctgt tgtctaccct cctgagatca agcaagagaa 240  
 tgcccttcat ttgcgagg ataccagcta tgctatggca cgtgccctca atgtcttgcc 300  
 aacttctat tcatatggtg att 323

<210> 113  
 <211> 312  
 <212> DNA  
 <213> Zea mays

<400> 113  
 cgataaggcc cttttcgaag agcttctacc gtcggatcaa cagattcttg gccgagctgc 60  
 tgtggcttca gcttgtctgg gtggtggact ggtgggcagg tgtaaggta caactgcatg 120  
 cagatgagga aacttacaga tcaatgggta aagagcatgc actcatcata tcaaatcatc 180  
 ggagtgatat tgattggctc attggatgga tattggccca gcgttcaggg tgccttgga 240  
 gtacacttgc tgtcatgaag aagtcaccca agttccttcc agttattggc tggatcaatg 300  
 ggtttgcaga gt 312

<210> 114  
 <211> 279  
 <212> DNA  
 <213> Zea mays

<400> 114  
 agtggggctc ccaaagggtg aaagacttcc ctagaccatt ttggctagct ctttttgttg 60  
 aggttactcg ctttactcca gcaaagcttc tcgcagctca ggagtatgcg gcttcccagg 120  
 gcttaccagc tcctagaaat gtacttatc cacgtaccaa gggatttcta tctgccgtaa 180  
 gtattatgcg agattttgtt ccagccattt acgatacaac tgtaatatgt cctaaagatt 240  
 cccctcaacc aacaatgctg cggattttga aagggaat 279

<210> 115  
 <211> 304  
 <212> DNA  
 <213> Zea mays

<400> 115  
 cgtcaacgcc atccaggccg tcctatttgt gacgataagg cccttttcga agagcttcta 60  
 ccgtcggatc aacagattct tggccgagct gctgtggctt cagcttctct ggggtgggga 120  
 ctggtgggca ggtgttaagg tacaactgca tgcagatgag gaaacttaca gatcaatggg 180  
 taaagagcat gactcatca tatcaaatca tcggagtgat attgattggc tcatggatgg 240  
 atattggccc agcgttcagg gtgccttgga agtacattgc tgtcatgaag aagtcaccca 300  
 agtt 304



<210> 116  
<211> 259  
<212> DNA  
<213> Zea mays

<400> 116  
cttctctctg tccggcctca tegtcaacgc catccaggcc gtcctatttg tgacgataag 60  
gccnttttcg aagagcttct aacgtcggat caacagattc ntggccgagc tgctgtggct 120  
tcagcttgct tgggtgggtg acnggtgggc aggtgttaag gtacaactgc atgcngatga 180  
ggaaacttac agatcnatgg gtanagagca tgcactcatc atatcaaate atcgagtgta 240  
tattgattgg cncattgga 259

<210> 117  
<211> 235  
<212> DNA  
<213> Zea mays

<400> 117  
attccacgta ccaagggatt tgtatctgct gtaagtatta tgcgagattt tgttccagcc 60  
atztatgata caactgtaat agttcctaaa gattcccctc aaccaacaat gctgcggatt 120  
ttgaaagggc aatcatcagt gatacatgtc cgcatgaaac gtcattgcaat gaggtagatg 180  
ccaaaatcag atgaggatgt ttcaaaatgg tgtaaagaca tttttgtggc aaagg 235

<210> 118  
<211> 282  
<212> DNA  
<213> Zea mays

<400> 118  
tgagatgcc aatcagatg atgacgtttc aaaatgggtg aaagacattt ttgtgacaaa 60  
ggatgcctta ctggacaaaac atttggcaac aggcactttc gatgaggaga ttagacctat 120  
cgcccgccca gtgaaatcat tgctgggtgac cctgttttgg tcgtgcctgc tgttgttttg 180  
tgccatcgag ttcttcaagt ggacgcagct cctatcgaca tggagaggag tggcattcac 240  
tgccgcagga tggcgctcgt gacaggggtc atgcacgtct tc 282

<210> 119  
<211> 166  
<212> DNA  
<213> Zea mays

<400> 119  
ctggtgggca ggcgttaagg tacaactaca tgcggatgag gacacttacc gatcaatggg 60  
taaagagcat gcactcgtca tatcaaatca tcgaagtgat attgattggc ttattggatg 120  
gatattggcc cagcgtcag ggtgccttgg aagtacgtc gctgtc 166

<210> 120  
<211> 234  
<212> DNA  
<213> Zea mays

<400> 120  
agtcancaaa gntccttcca gtcattggct ggtcaatgtg gtttgcagag tacctctttt 60  
nggagaggag ctgggccaag gatgaaaaga cactaaagtg ggtctctcaa aggttgaaag 120  
acttccctag accatttngg ctagctcttn ttgtngagg gnantcgtt tactccagca 180  
angnttntng aggnnncagn agnnncgggn ttcccanggg ttaacagncc cana 234

<210> 121  
<211> 210  
<212> DNA  
<213> Zea mays

<400> 121  
gtgagatgcn aaaatcagat gatgacgttt caaaatgggt taaagacatt tttgtggaca 60  
aaggatgcct tactggacaa acatttggca acaggcactt tcgatgagga gattagacct 120  
atcgcccgcc cagtgaatc atngctgggt accctgtntt ggtcgtgcct gctgttgttt 180  
ggtgccatcg agntcttcaa gtggacgcag 210

<210> 122  
<211> 274  
<212> DNA

<213> Zea mays

<400> 122

```
acncccgaaat cgcgcgcgcg cgcnccgctc tcgtcgccgg cggaggcgcc cgcnaccgcc 60
cacagcagcc tatcgccgga gaaggaaacgc cgcggggagc tttccacng ccatctcccg 120
tctgaccctt cgcgagatcgn aagcggcgcc catggcgatc cgcgtcgtgc tcgtcgtgct 180
cccgtcggc ctctctcttc tctgtccgg cctcatcgtc aacaccatcc aggccatcct 240
atttgtgaca ataaggccct tttccaagag cttg 274
```

<210> 123

<211> 305

<212> DNA

<213> Zea mays

<400> 123

```
ttgcactgag gaaaggccat tagggatata tcaagtacat acataagagc agcttgatga 60
agttgcctat ttttagctgg gcatttcaca tttttgagtt tatcccgga gaacggaaat 120
gggagattga tgaagcaatt attcagaaca agctatcaaa atttaagaac ccgagagatc 180
ctatctgggt ggccggttttt cctgaaggca cggattatac tgagaagaaa tgcacatga 240
gtcaaggata tgcttcagaa catggcttgc ctatgctaga acatgtcctc cttccaaaga 300
caagg 305
```

<210> 124

<211> 279

<212> DNA

<213> Zea mays

<400> 124

```
ccagattttc tggacaatgt gtatggcggt gatccttctg aagtccacat ccacgtcaga 60
atggttcagc tccatcacat cccacaaca gaagacaaga taacagaatg gatggncgag 120
aggtttaggc agaaggacca gctcctggca gatttcttca tgaaggggca tttcctgatg 180
aaaggaactg aaaggagatc tgctgacgcc gagtgcctgg caaactttct taaccagtag 240
tatgtctgac ggcnatctg gtttgtaact aaactcttt 279
```

<210> 125

<211> 219

<212> DNA

<213> Zea mays

<400> 125

```
agattttntg gacaatgtgt atggngttga tccttntgaa gtnacacatc acgtnagaat 60
ggttcagctc catcacatcc ccacaacagn agacaagata acagaangga tggtagagag 120
gttttaggcag aaggaccagc tcctggcaga tttcttcag aaggggcact ttcctgatga 180
aggaactgaa ggagatctgt cgacgccgaa gtgcctggc 219
```

<210> 126

<211> 293

<212> DNA

<213> Zea mays

<400> 126

```
taccatagat gctgtgtacg acatcacgat cgcntacaaa caccggcngc ngacatttct 60
ngacaacgtc tacngcgtgg ntccttcgga agtccacatc cacatcanca gcatccaggt 120
ctccgacata nccgctccg aaaaacgggg tggctggcng gntnngtgga gcggttcaag 180
gcntnganna acgagctngc tggtcggggc tttctaccgc ggctggggcc aatttcnccc 240
cgaacgaaa ggaaaaagg gaaccgaagg ggggaacctg ttngaacggg ncc 293
```

<210> 127

<211> 6

<212> PRT

<213> conserved sequence

<400> 127

Val Xaa Asn His Xaa Ser

1

5

<210> 128

<211> 6

<212> PRT

<213> conserved sequence

<400> 128

Val Thr Tyr Ser Xaa Ser  
1 5

<210> 129

<211> 7

<212> PRT

<213> conserved sequence

<400> 129

Val Xaa Leu Thr Arg Xaa Arg  
1 5

<210> 130

<211> 5

<212> PRT

<213> conserved sequence

<400> 130

Cys Pro Glu Gly Thr  
1 5

<210> 131

<211> 5

<212> PRT

<213> conserved sequence

<400> 131

Ile Val Pro Val Ala  
1 5

<210> 132

<211> 7

<212> PRT

<213> conserved sequence

<400> 132

Leu Xaa Xaa Gly Asp Leu Val  
1 5

<210> 133

<211> 6

<212> PRT

<213> conserved sequence

<400> 133

Phe Xaa Xaa Gly Ala Phe  
1 5

<210> 134

<211> 6

<212> PRT

<213> Synthetic Oligonucleotide

<400> 134

Val Ala Asn Xaa Xaa Gln  
1 5

<210> 135

<211> 30

<212> DNA

<213> Synthetic Oligonucleotide

<400> 135  
ccatccgctt caagggaaacg acacccatca 30

<210> 136  
<211> 31  
<212> DNA  
<213> Synthetic Oligonucleotide

<400> 136  
tcctgtctt gcttgatgaa cttaaagctt g 31

<210> 137  
<211> 30  
<212> DNA  
<213> Synthetic Oligonucleotide

<400> 137  
acagcaggag tgtctgatga tggcagattc 30

<210> 138  
<211> 30  
<212> DNA  
<213> Synthetic Oligonucleotide

<400> 138  
actggagttc cagccaaaaa tgacacctgc 30

<210> 139  
<211> 30  
<212> DNA  
<213> Synthetic Oligonucleotide

<400> 139  
gatacacctt tgaaatcagg cgattttgct 30

<210> 140  
<211> 30  
<212> DNA  
<213> Synthetic Oligonucleotide

<400> 140  
ttgcaaattc aattcctgtt tcaccggggc 30

<210> 141  
<211> 30  
<212> DNA  
<213> Synthetic Oligonucleotide

<400> 141  
gttttctgct attccagaag gcgtcaacaa 30

<210> 142  
<211> 32  
<212> DNA  
<213> Synthetic Oligonucleotide

<400> 142  
cattgaagat ccgtccgtga agttncctta cc 32

<210> 143  
<211> 30  
<212> DNA  
<213> Synthetic Oligonucleotide

<400> 143  
tcgagctgtg atcgatgatt ggctgtgaag 30

<210> 144

<211> 30  
<212> DNA  
<213> Synthetic Oligonucleotide

<400> 144  
gtctcttcaa aaacacacac acacgtctct 30

<210> 145  
<211> 30  
<212> DNA  
<213> Synthetic Oligonucleotide

<400> 145  
gtctcttcaa aaacacacac acacgtctct 30

<210> 146  
<211> 30  
<212> DNA  
<213> Synthetic Oligonucleotide

<400> 146  
gtagagagcc ttacttgctt cggtttagtc 30

<210> 147  
<211> 30  
<212> DNA  
<213> Synthetic Oligonucleotide

<400> 147  
acgtcatcgt acctggtgct attgactcac 30

<210> 148  
<211> 30  
<212> DNA  
<213> Synthetic Oligonucleotide

<400> 148  
acttttccat tgtcaggac tcttcgacac 30

<210> 149  
<211> 30  
<212> DNA  
<213> Synthetic Oligonucleotide

<400> 149  
acggtgtagg aagggaagg attcaaaagg 30

<210> 150  
<211> 30  
<212> DNA  
<213> Synthetic Oligonucleotide

<400> 150  
gcgatgaact acagagtcgg attcttcctc 30

<210> 151  
<211> 30  
<212> DNA  
<213> Synthetic Oligonucleotide

<400> 151  
ccggtttacg agattacgctt cttgaaccag 30

<210> 152  
<211> 30  
<212> DNA  
<213> Synthetic Oligonucleotide

<400> 152  
caatggagac aaggctcgaa agtgctaacc 30

<210> 153  
<211> 30  
<212> DNA  
<213> Synthetic Oligonucleotide

<400> 153  
attctctgaa catagtctgc cacggtcatg 30

<210> 154  
<211> 30  
<212> DNA  
<213> Synthetic Oligonucleotide

<400> 154  
gaaatccaac gccttcccaa tatcactctg 30

<210> 155  
<211> 30  
<212> DNA  
<213> Synthetic Oligonucleotide

<400> 155  
cttcaacttt ccatcaggat cttggcacgt 30

<210> 156  
<211> 30  
<212> DNA  
<213> Synthetic Oligonucleotide

<400> 156  
accacttggt agagacctta cctgcttagg 30

<210> 157  
<211> 30  
<212> DNA  
<213> Synthetic Oligonucleotide

<400> 157  
tcctacctac accatccaat ttctcgaccc 30

<210> 158  
<211> 30  
<212> DNA  
<213> Synthetic Oligonucleotide

<400> 158  
ctgcgtcaag tgagcaactc agttcttgca 30

<210> 159  
<211> 30  
<212> DNA  
<213> Synthetic Oligonucleotide

<400> 159  
tggaagcag cacgttggtc agtatcgga 30

<210> 160  
<211> 30  
<212> DNA  
<213> Synthetic Oligonucleotide

<400> 160  
tagcctctgt gtaatctgtg ccctcgggga 30

<210> 161  
<211> 1702  
<212> DNA  
<213> *Simmondsia chinensis*

&lt;400&gt; 161

```

gaattctagc ctctctctctc ctgcaattct atttgctttc tacgatcttt cctctctctc 60
ctctaaaacc ttaaaattgg aatggaatcg tttaaaaata tgatcttttt gtaattgaat 120
tagtataaatt atatctgggt aatcttgaat ttgttggtga ggccatgggg atcccagctg 180
cggctgtgat tgtaccgctt ggcttgcctc tcttctctc tggctctctc atcaacttca 240
ttcaggcaat ttgttttggt ctctgtcggc cactgtcaaa gntacatac agaaggatta 300
acagggtgct ggtggaattg ttgtggcttg agctgatatg gctcgtatag tgggtgggcaa 360
gtgttaagat caagttgttc acagatcctg atacctttcg gctaattggg aaagagcatg 420
cacttgtgat atcaaaccac agaagtgata ttgattggct tggttggatg gtgttggccc 480
agagatcagg ctgcctggga agcacactgg ctgtcatgaa gaaatcatca aagtttctcc 540
cggtcatagg ttggtctatg tggttttctg agtacctttt tcttgagaga agtggggcca 600
aggatgaaag cacattgaag ttaggtcttc aacgcctcaa ggactaccct ctgcctttct 660
ggttggctct tttcgtagaa ggaacacgat ttacccaagc taaactttta gcagctcaag 720
aatatgctac ttcaatggga ttgccagttc ctagaataac tttgatccct cgtactaagg 780
gatttggttc agccgtgagc catatgcgtt cgtttgtccc ggccatata gatgtaacgg 840
tggccatccc taaatcttct tcgcagccta caatgctcag acttttcaaa ggccagccat 900
ccacggttca tgtacacatc aagcgcgct cgatgaaaga tctccctgaa gcagcagatg 960
atgttgacac atggtgtcga gacacattcg tcgcaaagga tgcactcctg gacaagcata
1020
atgtagatga cactttcgga gatgagtatc tgcaggacac tggccggcct ttgaaatctc
1080
tctttgtagc agtctcttgg gcattgattc tcctcctggg aggtttgaaa ttcctacgat
1140
ggtcgtccct tctatcatca tgggaagggg tcgccttctc agccgcatgc cttgtgctcg
1200
tcaccattct tatgcagatc ttaatccaat tttctcaatc cgagcgctcg actcctgcta
1260
aggtagcccc aggaaagccc aagaacatgg tatcagaacc cagggaaacg caacgacata
1320
agcagcacta aaagtatata tggaccccaa ctaagaagat tcagacgcaa gccacagttg
1380
attcaactgt tcagaatgtc aaatatagtt tgagaaacaa aagatcaaga ttagctgatg
1440
aagagcctaa tgaacctaca tacttggatc tgtcgtcgcc accgtctgct gctagctcgt
1500
tatcagaatt cgtgattccg ggaccgatcc cggatcttag ccttctatgc atggattatg
1560
atagtatctt aaatttcttt aatgatgtac cggaattata atgttagtta attaggggga
1620
tgagcattgt ttgggtttat atcgtggtaa atccttgat tgtttataag atttgaagaa
1680
aattcgattc gagtgtctcg aa
1702

```

&lt;210&gt; 162

&lt;211&gt; 387

&lt;212&gt; PRT

<213> *Simmondsia chinensis*

&lt;400&gt; 162

```

Met Gly Ile Pro Ala Ala Val Ile Val Pro Leu Gly Leu Leu Phe
  1           5           10           15

Phe Phe Ser Gly Leu Phe Ile Asn Phe Ile Gln Ala Ile Cys Phe Val
          20           25           30

Leu Val Arg Pro Leu Ser Lys Thr Tyr Arg Arg Ile Asn Arg Val Leu
          35           40           45

Val Glu Leu Leu Trp Leu Glu Leu Ile Trp Leu Val Asp Trp Trp Ala
  50           55           60

Ser Val Lys Ile Lys Leu Phe Thr Asp Pro Asp Thr Phe Arg Leu Met
  65           70           75           80

Gly Lys Glu His Ala Leu Val Ile Ser Asn His Arg Ser Asp Ile Asp
          85           90           95

Trp Leu Val Gly Trp Val Leu Ala Gln Arg Ser Gly Cys Leu Gly Ser
 100           105           110

```

Thr Leu Ala Val Met Lys Lys Ser Ser Lys Phe Leu Pro Val Ile Gly  
 115 120 125  
 Trp Ser Met Trp Phe Ser Glu Tyr Leu Phe Leu Glu Arg Ser Trp Ala  
 130 135 140  
 Lys Asp Glu Ser Thr Leu Lys Leu Gly Leu Gln Arg Leu Lys Asp Tyr  
 145 150 155 160  
 Pro Leu Pro Phe Trp Leu Ala Leu Phe Val Glu Gly Thr Arg Phe Thr  
 165 170 175  
 Gln Ala Lys Leu Leu Ala Ala Gln Glu Tyr Ala Thr Ser Met Gly Leu  
 180 185 190  
 Pro Val Pro Arg Asn Thr Leu Ile Pro Arg Thr Lys Gly Phe Val Ser  
 195 200 205  
 Ala Val Ser His Met Arg Ser Phe Val Pro Ala Ile Tyr Asp Val Thr  
 210 215 220  
 Val Ala Ile Pro Lys Ser Ser Ser Gln Pro Thr Met Leu Arg Leu Phe  
 225 230 235 240  
 Lys Gly Gln Pro Ser Thr Val His Val His Ile Lys Arg Arg Ser Met  
 245 250 255  
 Lys Asp Leu Pro Glu Ala Ala Asp Asp Val Ala Gln Trp Cys Arg Asp  
 260 265 270  
 Thr Phe Val Ala Lys Asp Ala Leu Leu Asp Lys His Asn Val Asp Asp  
 275 280 285  
 Thr Phe Gly Asp Glu Tyr Leu Gln Asp Thr Gly Arg Pro Leu Lys Ser  
 290 295 300  
 Leu Phe Val Ala Val Ser Trp Ala Leu Ile Leu Ile Leu Gly Gly Leu  
 305 310 315 320  
 Lys Phe Leu Arg Trp Ser Ser Leu Leu Ser Ser Trp Lys Gly Val Ala  
 325 330 335  
 Phe Ser Ala Ala Cys Leu Val Leu Val Thr Ile Leu Met Gln Ile Leu  
 340 345 350  
 Ile Gln Phe Ser Gln Ser Glu Arg Ser Thr Pro Ala Lys Val Ala Pro  
 355 360 365  
 Gly Lys Pro Lys Asn Met Val Ser Glu Pro Thr Glu Thr Gln Arg His  
 370 375 380  
 Lys Gln His  
 385

&lt;210&gt; 163

&lt;211&gt; 43

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence:Synthetic  
Oligonucleotide

&lt;400&gt; 163

aagcttgcat gcgtcgacac aatggttcac gcgaccaagt cag

43

&lt;210&gt; 164

&lt;211&gt; 35



<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:Synthetic  
Oligonucleotide

<400> 164  
ggtaccgtcg actcacttct tgggtgtggt gatag 35

<210> 165  
<211> 44  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:Synthetic  
Oligonucleotide

<400> 165  
ggatccgcgg ccgcacaatg acgagcttta ctacttcct tcat 44

<210> 166  
<211> 38  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:Synthetic  
Oligonucleotide

<400> 166  
ggatcccctg caggtagag atccattgat tctgcaat 38

<210> 167  
<211> 38  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:Synthetic  
Oligonucleotide

<400> 167  
ggatccgcgg ccgcataatg gaatcagagc tcaaagat 38

<210> 168  
<211> 38  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:Synthetic  
Oligonucleotide

<400> 168  
ggatcccctg caggtcattc ttctttctga tggaaatc 38

<210> 169  
<211> 41  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:Synthetic  
Oligonucleotide

<400> 169  
ggatccgcgg ccgcacaatg actcggtcac aagatgtttc a 41

<210> 170  
<211> 38  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:Synthetic  
Oligonucleotide

<400> 170  
ggatcccctg caggtcactt ctcttccaat ctgcccag

38

<210> 171  
<211> 46  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:Synthetic  
Oligonucleotide

<400> 171  
ggatccgcgg cgcacaatg tccggttaata agatctcgac tcttca

46

<210> 172  
<211> 46  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:Synthetic  
Oligonucleotide

<400> 172  
ggatcccctg caggttattt tttcttgaca actccgttat taccgg

46

<210> 173  
<211> 39  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:Synthetic  
Oligonucleotide

<400> 173  
atatccgcgg cgcacaatg gttatggagc aagctggaa

39

<210> 174  
<211> 38  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:Synthetic  
Oligonucleotide

<400> 174  
ggatcccctg caggtcaatg gagacaaggc tcgaaagt

38

<210> 175  
<211> 42  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:Synthetic  
Oligonucleotide

<400> 175

ggatccgcgg cgcacaatg tccgccaaga tttcaatatt cc

42

<210> 176

<211> 38

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:Synthetic  
Oligonucleotide

<400> 176

ggatcccctg caggtaatt tttcttaact actccatt

38

<210> 177

<211> 42

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:Synthetic  
Oligonucleotide

<400> 177

ggatccgcgg cgcacaatg ggagctcagg agaaacggcg cc

42

<210> 178

<211> 38

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:Synthetic  
Oligonucleotide

<400> 178

ggatcccctg caggtcacgt cttctccttc ttcaccgg

38

<210> 179

<211> 44

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:Synthetic  
Oligonucleotide

<400> 179

ggatccgcgg cgcacaatg gcggatcctg atctgtcttc tcct

44

<210> 180

<211> 44

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:Synthetic  
Oligonucleotide

<400> 180

ggatcccctg caggttatgt tggggccaag tcagggtgcaa agat

44

<210> 181

<211> 44

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:Synthetic  
Oligonucleotide

<400> 181  
ggatccgcgg cgcgcaaatg gaaaaaaga gtgtaccaa ttct 44

<210> 182  
<211> 46  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:Synthetic  
Oligonucleotide

<400> 182  
ggatccctg cagggtatgt gtttactaat ttgagggaat tttttg 46

<210> 183  
<211> 36  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:Synthetic  
Oligonucleotide

<400> 183  
tcgacctgca ggaagcttaa ggatggtgat tgctgc 36

<210> 184  
<211> 31  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:Synthetic  
Oligonucleotide

<400> 184  
ggatccgcgg ccgcttactt ctccttctcc g 31

<210> 185  
<211> 39  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:Synthetic  
Oligonucleotide

<400> 185  
ggatccgcgg ccgcacaatg tcttttaggg atgtcctag 39

<210> 186  
<211> 41  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:Synthetic  
Oligonucleotide

<400> 186  
ggatccctg cagggtcaatc atccttacc tttgggttac c 41

<210> 187  
<211> 60  
<212> DNA  
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:Synthetic  
Oligonucleotide

<400> 187  
atgtctttta gggatgtcct agaaagagga gatgaatttt ctgtgcggta tttcacaccg 60

<210> 188  
<211> 60  
<212> DNA  
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:Synthetic  
Oligonucleotide

<400> 188  
tcaatcatcc ttaccctttg gtttaccctc tggaggcaga agattgtact gagagtgcac 60

<210> 189  
<211> 44  
<212> DNA  
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:Synthetic  
Oligonucleotide

<400> 189  
ggatccgcgg ccgcacaatg aagcattccc aaaaataccg tagg 44

<210> 190  
<211> 41  
<212> DNA  
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:Synthetic  
Oligonucleotide

<400> 190  
ggatcccctg caggatcaatg attttttttc atcacaaata c 41

<210> 191  
<211> 60  
<212> DNA  
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:Synthetic  
Oligonucleotide

<400> 191  
atgaagcatt cccaaaaata ccgtaggtat ggaatttatg ctgtgcggta tttcacaccg 60

<210> 192  
<211> 60  
<212> DNA  
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:Synthetic  
Oligonucleotide

<400> 192  
tcaatgattt tttttcatca caaatacaag aataagaaaa agattgtact gagagtgcac 60

<210> 193  
<211> 43  
<212> DNA  
<213> Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence:Synthetic  
Oligonucleotide

&lt;400&gt; 193

ggatccgcgg ccgcacaatg ggttttgttg atttcttcga aac

43

&lt;210&gt; 194

&lt;211&gt; 45

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence:Synthetic  
Oligonucleotide

&lt;400&gt; 194

ggatcccttg cagggtatttt ggtctcaatt ttaatatttt ttg

45

&lt;210&gt; 195

&lt;211&gt; 60

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence:Synthetic  
Oligonucleotide

&lt;400&gt; 195

atggggttttg ttgatttctt cgaaacatat atgggtcgggt ctgtgcggta tttcacaccg 60

&lt;210&gt; 196

&lt;211&gt; 60

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence:Synthetic  
Oligonucleotide

&lt;400&gt; 196

ttatttggtc tcaattttta ttttttttg caaggactcg agattgtact gagagtgcac 60

&lt;210&gt; 197

&lt;211&gt; 44

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence:Synthetic  
Oligonucleotide

&lt;400&gt; 197

ggatccgcgg ccgcacaatg gaaaagtaca ccaattggag agac

44

&lt;210&gt; 198

&lt;211&gt; 42

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence:Synthetic  
Oligonucleotide

&lt;400&gt; 198

ggatcccttg caggctactt cctcttttta cgttgatcgc tg

42

&lt;210&gt; 199

&lt;211&gt; 60

<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:Synthetic  
Oligonucleotide

<400> 199  
atggaaaagt acaccaattg gagagacaat ggtacgggaa ctgtgcggtta tttcacaccg 60

<210> 200  
<211> 60  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:Synthetic  
Oligonucleotide

<400> 200  
ctacttcctc tttttacgtt gatcgctgat atattccttc agattgtact gagagtgcac 60

<210> 201  
<211> 41  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:Synthetic  
Oligonucleotide

<400> 201  
ggatcgcgcg cgcacaatg cctgcaccaa aactcacgga g 41

<210> 202  
<211> 38  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:Synthetic  
Oligonucleotide

<400> 202  
ggatcccctg caggctacgc atctccttct ttcccttc 38

<210> 203  
<211> 60  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:Synthetic  
Oligonucleotide

<400> 203  
atgcctgcac caaaactcac ggagaaatct gcctcttcca ctgtgcggtta tttcacaccg 60

<210> 204  
<211> 60  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:Synthetic  
Oligonucleotide

<400> 204  
ctacgcatct ccttctttcc cttcttcttc ttcttcctct agattgtact gagagtgcac 60

<210> 205  
<211> 46  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:Synthetic  
Oligonucleotide

<400> 205  
ggatccgcgg cgcacaaatg tctgtcccg ctgccgatca taacgc 46

<210> 206  
<211> 44  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:Synthetic  
Oligonucleotide

<400> 206  
ggatccctcg caggtcattc tttcttttcg tgttctcttt tctg 44

<210> 207  
<211> 60  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:Synthetic  
Oligonucleotide

<400> 207  
atgtctgctc cgcgtgccga tcataacgct gccaaaccta ctgtgcggta tttcacaccg 60

<210> 208  
<211> 60  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:Synthetic  
Oligonucleotide

<400> 208  
tcattctttc ttttcgtgtt ctcttttctg tcttaccagc agattgtact gagagtgcac 60

<210> 209  
<211> 49  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:Synthetic  
Oligonucleotide

<400> 209  
ggatccgcgg cgcacaaatg ctgcatcaaa aaatagctca taaagttcg 49

<210> 210  
<211> 49  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:Synthetic  
Oligonucleotide

<400> 210



ggatcccctg cagggtcaaaa aataaaacaa taaagtttat aaactaacc

49

&lt;210&gt; 211

&lt;211&gt; 60

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence:Synthetic  
Oligonucleotide

&lt;400&gt; 211

atgctgcac tcaaaaatagc tcataaagtt cgaaaagtcg ctgtgcggtta ttccacaccg 60

&lt;210&gt; 212

&lt;211&gt; 60

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence:Synthetic  
Oligonucleotide

&lt;400&gt; 212

tcaaaaaata aaacaataaa gtttataaac taaccaaatt agattgtact gagagtgcac 60

&lt;210&gt; 213

&lt;211&gt; 41

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence:Synthetic  
Oligonucleotide

&lt;400&gt; 213

ggatccgcgg ccgcacaatg agtgtgatag gtaggttctt g

41

&lt;210&gt; 214

&lt;211&gt; 41

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence:Synthetic  
Oligonucleotide

&lt;400&gt; 214

ggatcccctg cagggttaatg catctttttt acagatgaac c

41

&lt;210&gt; 215

&lt;211&gt; 60

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence:Synthetic  
Oligonucleotide

&lt;400&gt; 215

atgagtgtga taggtagggtt cttgtattac ttgaggtccg ctgtgcggtta ttccacaccg 60

&lt;210&gt; 216

&lt;211&gt; 60

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence:Synthetic  
Oligonucleotide

<400> 216  
 ttaatgcac ttttttacag atgaaccttc gttatgggta agattgtact gagagtgcac 60

<210> 217

<211> 381

<212> PRT

<213> Saccharomyces sp.

<220>

<400> 217

Met Ser Phe Arg Asp Val Leu Glu Arg Gly Asp Glu Phe Leu Glu Ala  
 1 5 10 15

Tyr Pro Arg Arg Ser Pro Leu Trp Arg Phe Leu Ser Tyr Ser Thr Ser  
 20 25 30

Leu Leu Thr Phe Gly Val Ser Lys Leu Leu Leu Phe Thr Cys Tyr Asn  
 35 40 45

Val Lys Leu Asn Gly Phe Glu Lys Leu Glu Thr Ala Leu Glu Arg Ser  
 50 55 60

Lys Arg Glu Asn Arg Gly Leu Met Thr Val Met Asn His Met Ser Met  
 65 70 75 80

Val Asp Asp Pro Leu Val Trp Ala Thr Leu Pro Tyr Lys Leu Phe Thr  
 85 90 95

Ser Leu Asp Asn Ile Arg Trp Ser Leu Gly Ala His Asn Ile Cys Phe  
 100 105 110

Gln Asn Lys Phe Leu Ala Asn Phe Phe Ser Leu Gly Gln Val Leu Ser  
 115 120 125

Thr Glu Arg Phe Gly Val Gly Pro Phe Gln Gly Ser Ile Asp Ala Ser  
 130 135 140

Ile Arg Leu Leu Ser Pro Asp Asp Thr Leu Asp Leu Glu Trp Thr Pro  
 145 150 155 160

His Ser Glu Val Ser Ser Ser Leu Lys Lys Ala Tyr Ser Pro Pro Ile  
 165 170 175

Ile Arg Ser Lys Pro Ser Trp Val His Val Tyr Pro Glu Gly Phe Val  
 180 185 190

Leu Gln Leu Tyr Pro Pro Phe Glu Asn Ser Met Arg Tyr Phe Lys Trp  
 195 200 205

Gly Ile Thr Arg Met Ile Leu Glu Ala Thr Lys Pro Pro Ile Val Val  
 210 215 220

Pro Ile Phe Ala Thr Gly Phe Glu Lys Ile Ala Ser Glu Ala Val Thr  
 225 230 235 240

Asp Ser Met Phe Arg Gln Ile Leu Pro Arg Asn Phe Gly Ser Glu Ile  
 245 250 255

Asn Val Thr Ile Gly Asp Pro Leu Asn Asp Asp Leu Ile Asp Arg Tyr  
 260 265 270

Arg Lys Glu Trp Thr His Leu Val Glu Lys Tyr Tyr Asp Pro Lys Asn  
 275 280 285

Pro Asn Asp Leu Ser Asp Glu Leu Lys Tyr Gly Lys Glu Ala Gln Asp  
 290 295 300

Leu Arg Ser Arg Leu Ala Ala Glu Leu Arg Ala His Val Ala Glu Ile

305				310				315				320				
Arg	Asn	Glu	Val	Arg	Lys	Leu	Pro	Arg	Glu	Asp	Pro	Arg	Phe	Lys	Ser	
				325					330					335		
Pro	Ser	Trp	Trp	Lys	Arg	Phe	Asn	Thr	Thr	Glu	Gly	Lys	Ser	Asp	Pro	
				340					345					350		
Asp	Val	Lys	Val	Ile	Gly	Glu	Asn	Trp	Ala	Ile	Arg	Arg	Met	Gln	Lys	
				355					360					365		
Phe	Leu	Pro	Pro	Glu	Gly	Lys	Pro	Lys	Gly	Lys	Asp	Asp				
				370					375					380		

```
<210> 218
<211> 396
<212> PRT
<213> Saccharomyces sp.
```

<220>

<400>	218																
Met	Lys	His	Ser	Gln	Lys	Tyr	Arg	Arg	Tyr	Gly	Ile	Tyr	Glu	Lys	Thr		
1				5					10					15			
Gly	Asn	Pro	Phe	Ile	Lys	Gly	Leu	Gln	Arg	Leu	Leu	Ile	Ala	Cys	Leu		
			20					25					30				
Phe	Ile	Ser	Gly	Ser	Leu	Ser	Ile	Val	Val	Phe	Gln	Ile	Cys	Leu	Gln		
		35					40					45					
Val	Leu	Leu	Pro	Trp	Ser	Lys	Ile	Arg	Phe	Gln	Asn	Gly	Ile	Asn	Gln		
	50					55					60						
Ser	Lys	Lys	Ala	Phe	Ile	Val	Leu	Leu	Cys	Met	Ile	Leu	Asn	Met	Val		
65					70					75					80		
Ala	Pro	Ser	Ser	Leu	Asn	Val	Thr	Phe	Glu	Thr	Ser	Arg	Pro	Leu	Lys		
				85					90					95			
Asn	Ser	Ser	Asn	Ala	Lys	Pro	Cys	Phe	Arg	Phe	Lys	Asp	Arg	Ala	Ile		
			100					105					110				
Ile	Ile	Ala	Asn	His	Gln	Met	Tyr	Ala	Asp	Trp	Ile	Tyr	Leu	Trp	Trp		
		115					120					125					
Leu	Ser	Phe	Val	Ser	Asn	Leu	Gly	Gly	Asn	Val	Tyr	Ile	Ile	Leu	Lys		
	130					135					140						
Lys	Ala	Leu	Gln	Tyr	Ile	Pro	Leu	Leu	Gly	Phe	Gly	Met	Arg	Asn	Phe		
145					150					155					160		
Lys	Phe	Ile	Phe	Leu	Ser	Arg	Asn	Trp	Gln	Lys	Asp	Glu	Lys	Ala	Leu		
				165					170					175			
Thr	Asn	Ser	Leu	Val	Ser	Met	Asp	Leu	Asn	Ala	Arg	Cys	Lys	Gly	Pro		
			180					185					190				
Leu	Thr	Asn	Tyr	Lys	Ser	Cys	Tyr	Ser	Lys	Thr	Asn	Glu	Ser	Ile	Ala		
		195					200					205					
Ala	Tyr	Asn	Leu	Ile	Met	Phe	Pro	Glu	Gly	Thr	Asn	Leu	Ser	Leu	Lys		
	210					215					220						
Thr	Arg	Glu	Lys	Ser	Glu	Ala	Phe	Cys	Gln	Arg	Ala	His	Leu	Asp	His		
225					230					235					240		
Val	Gln	Leu	Arg	His	Leu	Leu	Leu	Pro	His	Ser	Lys	Gly	Leu	Lys	Phe		
				245					250					255			

Ala Val Glu Lys Leu Ala Pro Ser Leu Asp Ala Ile Tyr Asp Val Thr  
                   260                  265                  270

Ile Gly Tyr Ser Pro Ala Leu Arg Thr Glu Tyr Val Gly Thr Lys Phe  
                   275                  280                  285

Thr Leu Lys Lys Ile Phe Leu Met Gly Val Tyr Pro Glu Lys Val Asp  
           290                  295                  300

Phe Tyr Ile Arg Glu Phe Arg Val Asn Glu Ile Pro Leu Gln Asp Asp  
   305                  310                  315                  320

Glu Val Phe Phe Asn Trp Leu Leu Gly Val Trp Lys Glu Lys Asp Gln  
                   325                  330                  335

Leu Leu Glu Asp Tyr Tyr Asn Thr Gly Gln Phe Lys Ser Asn Ala Lys  
                   340                  345                  350

Asn Asp Asn Gln Ser Ile Val Val Thr Thr Gln Thr Thr Gly Phe Gln  
           355                  360                  365

His Glu Thr Leu Thr Pro Arg Ile Leu Ser Tyr Tyr Gly Phe Phe Ala  
           370                  375                  380

Phe Leu Ile Leu Val Phe Val Met Lys Lys Asn His  
   385                  390                  395

&lt;210&gt; 219

&lt;211&gt; 479

&lt;212&gt; PRT

&lt;213&gt; Saccharomyces sp.

&lt;220&gt;

&lt;400&gt; 219

Met Gly Phe Val Asp Phe Phe Glu Thr Tyr Met Val Gly Ser Arg Val  
   1                  5                  10                  15

Gln Phe Lys Gln Leu Asp Ile Ser Asp Trp Leu Ser Leu Thr Pro Arg  
           20                  25                  30

Leu Leu Ile Leu Phe Gly Tyr Phe Tyr Leu His Ser Phe Phe Thr Ala  
           35                  40                  45

Ile Asn Gln Phe Leu Gln Phe Ile Asn Thr Asn Ser Phe Cys Leu Arg  
           50                  55                  60

Leu His Leu Leu Tyr Asp Arg Phe Trp Ser His Val Pro Ile Ile Gly  
   65                  70                  75                  80

Glu Tyr Lys Ile Arg Leu Leu Ser Arg Ala Leu Thr Tyr Ser Lys Leu  
                   85                  90                  95

Lys Ile Ile Pro Thr Leu Asp Lys Val Leu Glu Ala Ile Glu Ile Trp  
                   100                  105                  110

Phe Gln Leu His Leu Val Glu Met Thr Phe Glu Lys Lys Lys Asn Val  
           115                  120                  125

Gln Ile Phe Ile Thr Glu Gly Ser Asp Asp Leu Asn Phe Phe Lys Asp  
           130                  135                  140

Ser Lys Phe Gln Thr Thr Leu Met Ile Cys Asn His Arg Ser Val Asn  
   145                  150                  155                  160

Asp Tyr Thr Leu Ile Asn Tyr Leu Phe Leu Lys Ser Cys Pro Thr Lys  
                   165                  170                  175

Phe Tyr Thr Lys Trp Glu Phe Leu Gln Lys Leu Arg Lys Gly Glu Asp  
 180 185 190  
 Leu Ala Glu Trp Pro Gln Leu Lys Phe Leu Gly Trp Gly Lys Met Phe  
 195 200 205  
 Asn Phe Pro Arg Leu Asp Leu Leu Lys Asn Ile Phe Phe Lys Asp Glu  
 210 215 220  
 Thr Leu Ala Leu Ser Ser Asn Glu Leu Arg Asp Ile Leu Glu Arg Gln  
 225 230 235 240  
 Asn Asn Gln Ala Ile Thr Ile Phe Pro Glu Val Asn Ile Met Ser Leu  
 245 250 255  
 Glu Leu Ser Ile Ile Gln Arg Lys Leu His Gln Asp Phe Pro Phe Val  
 260 265 270  
 Ile Asn Phe Tyr Asn Leu Leu Tyr Pro Arg Phe Lys Asn Phe Thr Thr  
 275 280 285  
 Leu Met Ala Ala Phe Ser Ser Ile Lys Asn Ile Lys Arg Lys Lys Asn  
 290 295 300  
 Arg Asn Asn Ile Ile Lys Glu Ala Arg Tyr Leu Phe His Arg Glu Leu  
 305 310 315 320  
 Asp Lys Leu Val His Lys Ser Met Lys Met Glu Ser Ser Lys Val Ser  
 325 330 335  
 Asp Lys Thr Thr Pro Pro Met Ile Val Asp Asn Ser Tyr Leu Leu Thr  
 340 345 350  
 Lys Lys Glu Glu Ile Ser Ser Gly Lys Pro Lys Val Val Arg Ile Asn  
 355 360 365  
 Pro Tyr Ile Tyr Asp Val Thr Ile Ile Tyr Tyr Arg Val Lys Tyr Thr  
 370 375 380  
 Asp Ser Gly His Asp His Thr Asn Gly Asp Leu Arg Leu His Lys Gly  
 385 390 395 400  
 Tyr Gln Leu Glu Gln Ile Ser Pro Thr Ile Phe Glu Met Ile Gln Pro  
 405 410 415  
 Glu Met Glu Ser Glu Asn Asn Ile Lys Asp Lys Asp Pro Ile Val Val  
 420 425 430  
 Met Val Asn Val Lys Lys His Gln Ile Gln Pro Leu Leu Ala Tyr Asn  
 435 440 445  
 Asp Glu Ser Leu Glu Lys Trp Leu Glu Asn Arg Trp Ile Glu Lys Asp  
 450 455 460  
 Arg Leu Ile Glu Ser Leu Gln Lys Asn Ile Lys Ile Glu Thr Lys  
 465 470 475

&lt;210&gt; 220

&lt;211&gt; 300

&lt;212&gt; PRT

&lt;213&gt; Saccharomyces sp.

&lt;400&gt; 220

Met Glu Lys Tyr Thr Asn Trp Arg Asp Asn Gly Thr Gly Ile Ala Pro  
 1 5 10 15

Phe Leu Pro Asn Thr Ile Arg Lys Pro Ser Lys Val Met Thr Ala Cys  
 20 25 30

Leu Leu Gly Ile Leu Gly Val Lys Thr Ile Ile Met Leu Pro Leu Ile  
                   35                                  40                                  45  
 Met Leu Tyr Leu Leu Thr Gly Gln Asn Asn Leu Leu Gly Leu Ile Leu  
                   50                                  55                                  60  
 Lys Phe Thr Phe Ser Trp Lys Glu Glu Ile Thr Val Gln Gly Ile Lys  
                   65                                  70                                  75                                  80  
 Lys Arg Asp Val Arg Lys Ser Lys His Tyr Pro Gln Lys Gly Lys Leu  
                                   85                                  90                                  95  
 Tyr Ile Cys Asn Cys Thr Ser Pro Leu Asp Ala Phe Ser Val Val Leu  
                                   100                                  105                                  110  
 Leu Ala Gln Gly Pro Val Thr Leu Leu Val Pro Ser Asn Asp Ile Val  
                   115                                  120                                  125  
 Tyr Lys Val Ser Ile Arg Glu Phe Ile Asn Phe Ile Leu Ala Gly Gly  
                   130                                  135                                  140  
 Leu Asp Ile Lys Leu Tyr Gly His Glu Val Ala Glu Leu Ser Gln Leu  
                   145                                  150                                  155                                  160  
 Gly Asn Thr Val Asn Phe Met Phe Ala Glu Gly Thr Ser Cys Asn Gly  
                                   165                                  170                                  175  
 Lys Ser Val Leu Pro Phe Ser Ile Thr Gly Lys Lys Leu Lys Glu Phe  
                                   180                                  185                                  190  
 Ile Asp Pro Ser Ile Thr Thr Met Asn Pro Ala Met Ala Lys Thr Lys  
                   195                                  200                                  205  
 Lys Phe Glu Leu Gln Thr Ile Gln Ile Lys Thr Asn Lys Thr Ala Ile  
                   210                                  215                                  220  
 Thr Thr Leu Pro Ile Ser Asn Met Glu Tyr Leu Ser Arg Phe Leu Asn  
                   225                                  230                                  235                                  240  
 Lys Gly Ile Asn Val Lys Cys Lys Ile Asn Glu Pro Gln Val Leu Ser  
                                   245                                  250                                  255  
 Asp Asn Leu Glu Glu Leu Arg Val Ala Leu Asn Gly Gly Asp Lys Tyr  
                                   260                                  265                                  270  
 Lys Leu Val Ser Arg Lys Leu Asp Val Glu Ser Lys Arg Asn Phe Val  
                   275                                  280                                  285  
 Lys Glu Tyr Ile Ser Asp Gln Arg Lys Lys Arg Lys  
                   290                                  295                                  300

&lt;210&gt; 221

&lt;211&gt; 759

&lt;212&gt; PRT

&lt;213&gt; Saccharomyces sp.

&lt;400&gt; 221

Met Pro Ala Pro Lys Leu Thr Glu Lys Phe Ala Ser Ser Lys Ser Thr  
                   1                                  5                                  10                                  15  
 Gln Lys Thr Thr Asn Tyr Ser Ser Ile Glu Ala Lys Ser Val Lys Thr  
                                   20                                  25                                  30  
 Ser Ala Asp Gln Ala Tyr Ile Tyr Gln Glu Pro Ser Ala Thr Lys Lys  
                   35                                  40                                  45  
 Ile Leu Tyr Ser Ile Ala Thr Trp Leu Leu Tyr Asn Ile Phe His Cys  
                   50                                  55                                  60

Phe Phe Arg Glu Ile Arg Gly Arg Gly Ser Phe Lys Val Pro Gln Gln  
 65 70 75 80  
 Gly Pro Val Ile Phe Val Ala Ala Pro His Ala Asn Gln Phe Val Asp  
 85 90 95  
 Pro Val Ile Leu Met Gly Glu Val Lys Lys Ser Val Asn Arg Arg Val  
 100 105 110  
 Ser Phe Leu Ile Ala Glu Ser Ser Leu Lys Gln Pro Pro Ile Gly Phe  
 115 120 125  
 Leu Ala Ser Phe Phe Met Ala Ile Gly Val Val Arg Pro Gln Asp Asn  
 130 135 140  
 Leu Lys Pro Ala Glu Gly Thr Ile Arg Val Asp Pro Thr Asp Tyr Lys  
 145 150 155 160  
 Arg Val Ile Gly His Asp Thr His Phe Leu Thr Asp Cys Met Pro Lys  
 165 170 175  
 Gly Leu Ile Gly Leu Pro Lys Ser Met Gly Phe Gly Glu Ile Gln Ser  
 180 185 190  
 Ile Glu Ser Asp Thr Ser Leu Thr Leu Arg Lys Glu Phe Lys Met Ala  
 195 200 205  
 Lys Pro Glu Ile Lys Thr Ala Leu Leu Thr Gly Thr Thr Tyr Lys Tyr  
 210 215 220  
 Ala Ala Lys Val Asp Gln Ser Cys Val Tyr His Arg Val Phe Glu His  
 225 230 235 240  
 Leu Ala His Asn Asn Cys Ile Gly Ile Phe Pro Glu Gly Gly Ser His  
 245 250 255  
 Asp Arg Thr Asn Leu Leu Pro Leu Lys Ala Gly Val Ala Ile Met Ala  
 260 265 270  
 Leu Gly Cys Met Asp Lys His Pro Asp Val Asn Val Lys Ile Val Pro  
 275 280 285  
 Cys Gly Met Asn Tyr Phe His Pro His Lys Phe Arg Ser Arg Ala Val  
 290 295 300  
 Val Glu Phe Gly Asp Pro Ile Glu Ile Pro Lys Glu Leu Val Ala Lys  
 305 310 315 320  
 Tyr His Asn Pro Glu Thr Asn Arg Asp Ala Val Lys Glu Leu Leu Asp  
 325 330 335  
 Thr Ile Ser Lys Gly Leu Gln Ser Val Thr Val Thr Cys Ser Asp Tyr  
 340 345 350  
 Glu Thr Leu Met Val Val Gln Thr Ile Arg Arg Leu Tyr Met Thr Gln  
 355 360 365  
 Phe Ser Thr Lys Leu Pro Leu Pro Leu Ile Val Glu Met Asn Arg Arg  
 370 375 380  
 Met Val Lys Gly Tyr Glu Phe Tyr Arg Asn Asp Pro Lys Ile Ala Asp  
 385 390 395 400  
 Leu Thr Lys Asp Ile Met Ala Tyr Asn Ala Ala Leu Arg His Tyr Asn  
 405 410 415  
 Leu Pro Asp His Leu Val Glu Glu Ala Lys Val Asn Phe Ala Lys Asn  
 420 425 430

Leu Gly Leu Val Phe Phe Arg Ser Ile Gly Leu Cys Ile Leu Phe Ser  
 435 440 445  
 Leu Ala Met Pro Gly Ile Ile Met Phe Ser Pro Val Phe Ile Leu Ala  
 450 455 460  
 Lys Arg Ile Ser Gln Glu Lys Ala Arg Thr Ala Leu Ser Lys Ser Thr  
 465 470 475 480  
 Val Lys Ile Lys Ala Asn Asp Val Ile Ala Thr Trp Lys Ile Leu Ile  
 485 490 495  
 Gly Met Gly Phe Ala Pro Leu Leu Tyr Ile Phe Trp Ser Val Leu Ile  
 500 505 510  
 Thr Tyr Tyr Leu Arg His Lys Pro Trp Asn Lys Ile Tyr Val Phe Ser  
 515 520 525  
 Gly Ser Tyr Ile Ser Cys Val Ile Val Thr Tyr Ser Ala Leu Ile Val  
 530 535 540  
 Gly Asp Ile Gly Met Asp Gly Phe Lys Ser Leu Arg Pro Leu Val Leu  
 545 550 555 560  
 Ser Leu Thr Ser Pro Lys Gly Leu Gln Lys Leu Gln Lys Asp Arg Arg  
 565 570 575  
 Asn Leu Ala Glu Arg Ile Ile Glu Val Val Asn Asn Phe Gly Ser Glu  
 580 585 590  
 Leu Phe Pro Asp Phe Asp Ser Ala Ala Leu Arg Glu Glu Phe Asp Val  
 595 600 605  
 Ile Asp Glu Glu Glu Glu Asp Arg Lys Thr Ser Glu Leu Asn Arg Arg  
 610 615 620  
 Lys Met Leu Arg Lys Gln Lys Ile Lys Arg Gln Glu Lys Asp Ser Ser  
 625 630 635 640  
 Ser Pro Ile Ile Ser Gln Arg Asp Asn His Asp Ala Tyr Glu His His  
 645 650 655  
 Asn Gln Asp Ser Asp Gly Val Ser Leu Val Asn Ser Asp Asn Ser Leu  
 660 665 670  
 Ser Asn Ile Pro Leu Phe Ser Ser Thr Phe His Arg Lys Ser Glu Ser  
 675 680 685  
 Ser Leu Ala Ser Thr Ser Val Ala Pro Ser Ser Ser Ser Glu Phe Glu  
 690 695 700  
 Val Glu Asn Glu Ile Leu Glu Glu Lys Asn Gly Leu Ala Ser Lys Ile  
 705 710 715 720  
 Ala Gln Ala Val Leu Asn Lys Arg Ile Gly Glu Asn Thr Ala Arg Glu  
 725 730 735  
 Glu Glu Glu Glu Glu Glu Glu Glu Glu Glu Glu Glu Glu Glu  
 740 745 750  
 Glu Gly Lys Glu Gly Asp Ala  
 755

&lt;210&gt; 222

&lt;211&gt; 743

&lt;212&gt; PRT

<213> *Saccharomyces* sp.

&lt;400&gt; 222



Met Ser Ala Pro Ala Ala Asp His Asn Ala Ala Lys Pro Ile Pro His  
 1 5 10 15  
 Val Pro Gln Ala Ser Arg Arg Tyr Lys Asn Ser Tyr Asn Gly Phe Val  
 20 25 30  
 Tyr Asn Ile His Thr Trp Leu Tyr Asp Val Ser Val Phe Leu Phe Asn  
 35 40 45  
 Ile Leu Phe Thr Ile Phe Phe Arg Glu Ile Lys Val Arg Gly Ala Tyr  
 50 55 60  
 Asn Val Pro Glu Val Gly Val Pro Thr Ile Leu Val Cys Ala Pro His  
 65 70 75 80  
 Ala Asn Gln Phe Ile Asp Pro Ala Leu Val Met Ser Gln Thr Arg Leu  
 85 90 95  
 Leu Lys Thr Ser Ala Gly Lys Ser Arg Ser Arg Met Pro Cys Phe Val  
 100 105 110  
 Thr Ala Glu Ser Ser Phe Lys Lys Arg Phe Ile Ser Phe Phe Gly His  
 115 120 125  
 Ala Met Gly Gly Ile Pro Val Pro Arg Ile Gln Asp Asn Leu Lys Pro  
 130 135 140  
 Val Asp Glu Asn Leu Glu Ile Tyr Ala Pro Asp Leu Lys Asn His Pro  
 145 150 155 160  
 Glu Ile Ile Lys Gly Arg Ser Lys Asn Pro Gln Thr Thr Pro Val Asn  
 165 170 175  
 Phe Thr Lys Arg Phe Ser Ala Lys Ser Leu Leu Gly Leu Pro Asp Tyr  
 180 185 190  
 Leu Ser Asn Ala Gln Ile Lys Glu Ile Pro Asp Asp Glu Thr Ile Ile  
 195 200 205  
 Leu Ser Ser Pro Phe Arg Thr Ser Lys Ser Lys Val Val Glu Leu Leu  
 210 215 220  
 Thr Asn Gly Thr Asn Phe Lys Tyr Ala Glu Lys Ile Asp Asn Thr Glu  
 225 230 235 240  
 Thr Phe Gln Ser Val Phe Asp His Leu His Thr Lys Gly Cys Val Gly  
 245 250 255  
 Ile Phe Pro Glu Gly Gly Ser His Asp Arg Pro Ser Leu Leu Pro Ile  
 260 265 270  
 Lys Ala Gly Val Ala Ile Met Ala Leu Gly Ala Val Ala Ala Asp Pro  
 275 280 285  
 Thr Met Lys Val Ala Val Val Pro Cys Gly Leu His Tyr Phe His Arg  
 290 295 300  
 Asn Lys Phe Arg Ser Arg Ala Val Leu Glu Tyr Gly Glu Pro Ile Val  
 305 310 315 320  
 Val Asp Gly Lys Tyr Gly Glu Met Tyr Lys Asp Ser Pro Arg Glu Thr  
 325 330 335  
 Val Ser Lys Leu Leu Lys Lys Ile Thr Asn Ser Leu Phe Ser Val Thr  
 340 345 350  
 Glu Asn Ala Pro Asp Tyr Asp Thr Leu Met Val Ile Gln Ala Ala Arg  
 355 360 365  
 Arg Leu Tyr Gln Pro Val Lys Val Arg Leu Pro Leu Pro Ala Ile Val



<210> 223  
 <211> 397  
 <212> PRT  
 <213> Saccharomyces sp.

<400> 223  
 Met Leu His Gln Lys Ile Ala His Lys Val Arg Lys Val Val Val Pro  
   1                  5                  10                  15  
 Gly Ile Ser Leu Leu Ile Phe Phe Gln Gly Cys Leu Ile Leu Leu Phe  
                   20                  25                  30  
 Leu Gln Leu Thr Tyr Lys Thr Leu Tyr Cys Arg Asn Asp Ile Arg Lys  
                   35                  40                  45  
 Gln Ile Gly Leu Asn Lys Thr Lys Arg Leu Phe Ile Val Leu Val Ser  
                   50                  55                  60  
 Ser Ile Leu His Val Val Ala Pro Ser Ala Val Arg Ile Thr Thr Glu  
   65                  70                  75                  80  
 Asn Ser Ser Val Pro Lys Gly Thr Phe Phe Leu Asp Leu Lys Lys Lys  
                   85                  90                  95  
 Arg Ile Leu Ser His Leu Lys Ser Asn Ser Val Ala Ile Cys Asn His  
                   100                  105                  110  
 Gln Ile Tyr Thr Asp Trp Ile Phe Leu Trp Trp Leu Ala Tyr Thr Ser  
                   115                  120                  125  
 Asn Leu Gly Ala Asn Val Phe Ile Ile Leu Lys Lys Ser Leu Ala Ser  
   130                  135                  140  
 Ile Pro Ile Leu Gly Phe Gly Met Arg Asn Tyr Asn Phe Ile Phe Met  
 145                  150                  155                  160  
 Ser Arg Lys Trp Ala Gln Asp Lys Ile Thr Leu Ser Asn Ser Leu Ala  
                   165                  170                  175  
 Gly Leu Asp Ser Asn Ala Arg Gly Ala Gly Ser Leu Ala Gly Lys Ser  
                   180                  185                  190  
 Pro Glu Arg Ile Thr Glu Glu Gly Glu Ser Ile Trp Asn Pro Glu Val  
                   195                  200                  205  
 Ile Asp Pro Lys Gln Ile His Trp Pro Tyr Asn Leu Ile Leu Phe Pro  
   210                  215                  220  
 Glu Gly Thr Asn Leu Ser Ala Asp Thr Arg Gln Lys Ser Ala Lys Tyr  
 225                  230                  235                  240  
 Ala Ala Lys Ile Gly Lys Lys Pro Phe Lys Asn Val Leu Leu Pro His  
                   245                  250                  255  
 Ser Thr Gly Leu Arg Tyr Ser Leu Gln Lys Leu Lys Pro Ser Ile Glu  
                   260                  265                  270  
 Ser Leu Tyr Asp Ile Thr Ile Gly Tyr Ser Gly Val Lys Gln Glu Glu  
                   275                  280                  285  
 Tyr Gly Glu Leu Ile Tyr Gly Leu Lys Ser Ile Phe Leu Glu Gly Lys  
   290                  295                  300  
 Tyr Pro Lys Leu Val Asp Ile His Ile Arg Ala Phe Asp Val Lys Asp  
 305                  310                  315                  320  
 Ile Pro Leu Glu Asp Glu Asn Glu Phe Ser Glu Trp Leu Tyr Lys Ile  
                   325                  330                  335

Trp Ser Glu Lys Asp Ala Leu Met Glu Arg Tyr Tyr Ser Thr Gly Ser  
                   340                  345                  350  
 Phe Val Ser Asp Pro Glu Thr Asn His Ser Val Thr Asp Ser Phe Lys  
                   355                  360                  365  
 Ile Asn Arg Ile Glu Leu Thr Glu Val Leu Ile Leu Pro Thr Leu Thr  
                   370                  375                  380  
 Ile Ile Trp Leu Val Tyr Lys Leu Tyr Cys Phe Ile Phe  
                   385                  390                  395

<210> 224  
 <211> 303  
 <212> PRT  
 <213> Saccharomyces sp.

<400> 224  
 Met Ser Val Ile Gly Arg Phe Leu Tyr Tyr Leu Arg Ser Val Leu Val  
   1                  5                  10                  15  
 Val Leu Ala Leu Ala Gly Cys Gly Phe Tyr Gly Val Ile Ala Ser Ile  
                   20                  25                  30  
 Leu Cys Thr Leu Ile Gly Lys Gln His Leu Ala Gln Trp Ile Thr Ala  
                   35                  40                  45  
 Arg Cys Phe Tyr His Val Met Lys Leu Met Leu Gly Leu Asp Val Lys  
                   50                  55                  60  
 Val Val Gly Glu Glu Asn Leu Ala Lys Lys Pro Tyr Ile Met Ile Ala  
   65                  70                  75                  80  
 Asn His Gln Ser Thr Leu Asp Ile Phe Met Leu Gly Arg Ile Phe Pro  
                   85                  90                  95  
 Pro Gly Cys Thr Val Thr Ala Lys Lys Ser Leu Lys Tyr Val Pro Phe  
                   100                  105                  110  
 Leu Gly Trp Phe Met Ala Leu Ser Gly Thr Tyr Phe Leu Asp Arg Ser  
                   115                  120                  125  
 Lys Arg Gln Glu Ala Ile Asp Thr Leu Asn Lys Gly Leu Glu Asn Val  
   130                  135                  140  
 Lys Lys Asn Lys Arg Ala Leu Trp Val Phe Pro Glu Gly Thr Arg Ser  
   145                  150                  155                  160  
 Tyr Thr Ser Glu Leu Thr Met Leu Pro Phe Lys Lys Gly Ala Phe His  
                   165                  170                  175  
 Leu Ala Gln Gln Gly Lys Ile Pro Ile Val Pro Val Val Val Ser Asn  
                   180                  185                  190  
 Thr Ser Thr Leu Val Ser Pro Lys Tyr Gly Val Phe Asn Arg Gly Cys  
                   195                  200                  205  
 Met Ile Val Arg Ile Leu Lys Pro Ile Ser Thr Glu Asn Leu Thr Lys  
   210                  215                  220  
 Asp Lys Ile Gly Glu Phe Ala Glu Lys Val Arg Asp Gln Met Val Asp  
   225                  230                  235                  240  
 Thr Leu Lys Glu Ile Gly Tyr Ser Pro Ala Ile Asn Asp Thr Thr Leu  
                   245                  250                  255  
 Pro Pro Gln Ala Ile Glu Tyr Ala Ala Leu Gln His Asp Lys Lys Val  
                   260                  265                  270

Asn Lys Lys Ile Lys Asn Glu Pro Val Pro Ser Val Ser Ile Ser Asn  
275 280 285

Asp Val Asn Thr His Asn Glu Gly Ser Ser Val Lys Lys Met His  
290 295 300

<210> 225

<211> 1146

<212> DNA

<213> *Saccharomyces* sp.

<400> 225

```
atgtctttta gggatgtcct agaaagagga gatgaatttt tagaagccta tcccagaaga 60
agcccccttt ggagatttct ttcatacagt acatcattac tgaccttcgg tgtatcaaaa 120
ctgcttcttt tcacatgcta taatgtcaaa ttgaatggtt ttgaaaaatt agaaactgcc 180
ttggaacggt ccaaaaggga aaatagaggg cttatgacgg tcatgaacca tatgagtatg 240
gtcgtatgat cgttagtttg ggcaacacta ccatataagt tatttacgtc ttgggacaac 300
ataagatggt ctttgggtgc acataatatt tgctttcaaa ataaatttct ggccaacttt 360
ttctcacttg gccaagtcct ttcaacagaa agatttgagg tgggcccatt tcaaggttct 420
atagatgctt caataagatt gttaagccct gacgacact tagacttgga atggaccctt 480
cactctgagg tctcttcttc gctaaaaaaa gcctactccc cgcccataat aaggtcgaag 540
ccatcttggg tccatgttta tccagaagga ttgtactac aattatatcc gccttttgaa 600
aattcgatga ggtattttta atggggtatt accagaatga tctagaagc aacaaagccg 660
cccattgtag taccaatatt tgctacaggg ttgaaaaaaa tagcatccga agcagtcaca 720
gattcaatgt ttagacaaat tctaccaaga aactttggct ctgaaataaa tgttaccata 780
ggggatcctt taaatgatga ttaatcgac aggtatagaa aagaatggac acatttggtt 840
gaaaaatact atgatcccaa aaatcctaac gacctctctg acgaattgaa atatggtaaa 900
gaggcgcaag atttaagaag cagattagcc gctgaactga gagcccatgt tgctgaaatt 960
agaaatgaag ttcgcaaatt accacgcgaa gaccctaggt tcaaatcccc ctcattggtg
1020
aagcgggttca acaccacgga aggtaaatcg gaccagatg ttaaagtcatt tggcgaaaaa
1080
tgggcaataa ggaggatgca aaagtttctg cctccagagg gtaaaccaaa gggttaaggat
1140
gattga
1146
```

<210> 226

<211> 1191

<212> DNA

<213> *Saccharomyces* sp.

<400> 226

```
atgaagcatt cccaaaaata ccgtaggtat ggaatttatg aaaagactgg taatcccttt 60
ataaaagggt tgcaaaggct gcttatcgct tgcattgttca ttccaggctc gctgagtatt 120
gtcgtttttc agatctgtct acagggtgct ctcccttgga gcaagattag atttcaaaat 180
gggtataaatc aaagtaagaa ggctttttatc gttttattat gcatgatctt gaacatgggtg 240
gctccctctt ctttgaatgt cacttttgaa acatcgcggc cattgaagaa ctcttctaac 300
gccaagccat gcttttagatt taaagacagg gctataataa ttgcaaatca tcaaatgtat 360
gcagactgga tttatctctg gtggctttcc tttgtttcaa atttggtggg taacgtttat 420
atcatcctga agaaagctct gcagtacata ccattactgg gatttggcat gcgaaatttt 480
aagtttatat ttttaagtag gaactggcaa aaggatgaga aagctttaac aaatagtttg 540
gtttctatgg acttaaacgc gaggtgcaag gggcccctta caaattataa gagttgttat 600
tccaagacaa atgaatccat tgccgcttat aatttaatca tgttccctga gggtagaaaa 660
ctaagcctca agacaagaga aaaaagcgag gcattctgtc aaagagcaca tttggaccat 720
gtccaattaa gacatttggt attaccgcac tctaaaggct tgaagtttgc agtagaaaaa 780
ctagctecta gtttagatgc tatctacgat gtcactattg gatattctcc cgcttgaga 840
acggaatagc tcggcaccaa attcaccttg aagaaaaat tcttaattgg tgcttatccg 900
gagaaagtag atttttatat tagggaattt agagttaatg agatcccttt gcaagatgac 960
gaagtttttt tcaattggtt actgggctg tggaagaaa aagatcaact gctagaagac
1020
tactacaaca caggccaatt taaaagtaat gctaaaaatg acaaccaatc catcgttggtt
1080
acgacacaaa cgactggatt tcagcacgaa acattgacac cccgtatcct ttcattattac
1140
gggttcttcg cttttcttat tcttgtattt gtgatgaaaa aaaatcattg a
1191
```

<210> 227  
<211> 1440  
<212> DNA  
<213> *Saccharomyces* sp.

<400> 227  
atgggtttttg ttgatttctt cgaacatat atgggtcggtt ctagggtcca gttcaaacag 60  
ttagatatatt ctgattgggt gagtctgacc ccaagggtgc ttattctttt tggctatttt 120  
taccttcatt cttttttttac tgcaatcaat caattcctac agttcattaa cacgaattcc 180  
ttctgtctta gactgcattt actatatgac agatttttgt cgcattgtgc cataataggt 240  
gagtacaaaa ttcggctgct ctcgagggca ctgacatata gtaaaactgaa aataatacca 300  
actttagaca aggtgctgga ggcgattgaa atttgggttc agctacattt agttgaaatg 360  
accttcgaaa aaaaaaaaaa cgtccaaatt ttcataaccg aggggaagtga tgacctaaac 420  
ttttttaaag atagcaaatt ccaaaccaca ttaatgatat gtaatcatcg atcagtgaat 480  
gactacacat tgattaatta cctttttctc aaaagttgtc ccaccaagtt ttatactaaa 540  
tggaattttc tacaaaagct gaggaagggg gaagatctag ctgaatggcc tcagttaaaa 600  
tttcttgggt ggggaaaaat gtttaacttt cctcgattgg atctactaaa gaacatattc 660  
ttcaaagatg aaacactcgc actctcatcg aatgagttaa gagatatttt agaaagacaa 720  
aacaatcaag ctattactat ttttcccgaa gtcaatatca tgagtttggg actatcaatt 780  
attcaaagaa aattacacca agattttccc tttgttataa acttctataa tttattatac 840  
ccaagattta aaaactttac cactttgatg gctgcttttt catcaattaa aaacatcaaa 900  
agaaagaaaa accgtaacaa tataatcaaa gagggccgat acctgtttca cagagaactt 960  
gacaaattag ttcacaagag catgaaaatg gagtcttcca aggtatccga taagacgacg  
1020  
ccgcccatag tcgtagataa ttcatactta cttacaaaaa aggaagaaat cagcagcggc  
1080  
aagcccaagg tggtagcaat caatccatac atatattgatg tcaccataat ttattaccga  
1140  
gtcaaataa ctgatagtgg gcattgatcat accaacggag atttgagact tcataaagggt  
1200  
tatcaattag agcaaatac tccgacaatc tttgagatga ttcaaccaga aatggagtct  
1260  
gaaaacaaca taaaggataa ggaccccatg gttgtgatgg taaatgtaaa aaagcatcaa  
1320  
attcaaccat tactcgcata caatgatgag agtttagaaa agtggcttga aaatagggtg  
1380  
atagaaaaag atagattaat cgagtccttg caaaaaata ttaaaattga gaccaaataa  
1440

<210> 228  
<211> 903  
<212> DNA  
<213> *Saccharomyces* sp.

<400> 228  
atggaaaagt acaccaattg gagagacaat ggtacgggaa tagctccatt tctaccaaac 60  
acaatcagga aacctagtaa ggtgatgaca gcgtgttgtg tgggtatcct aggggtgaaa 120  
gattatgataa tgctaccatt gattatgctg taccttctaa ctggccagaa caacttactg 180  
ggtttgatat tgaagtttac attcagttgg aaagaggaaa ttaccgtgca aggaatcaag 240  
aaacgtgacg taaggaaatc caagcattat ccacagaagg gcaagcttta tatttgcaat 300  
tgtacctcac ctttagatgc ttttccagtg gtgttattag ctcaaggggc tgttacgttg 360  
ttggtcccat ccaatgacat tgtatacaaa gtttccataa gagaattcat caacttcac 420  
ctcgcgggtg ggttagatat aaaactctat ggccacgagg tagcagagct atctcaattg 480  
ggcaataccg tgaattttat gtttgctgag ggtacctcat gtaatggtaa aagcgtctta 540  
ccgttttagta taaccgggaa aaaacttaaa gaattcatag acccttcaat aaccacaatg 600  
aaccgccgaa tggccaaaac taaaaaattt gaattgcaga ccatccaaat caaaactaat 660  
aaaactgcca tcaccacatt gcccatctcc aatatggagt atttatctag atttctgaac 720  
aagggcatta atgttaaag caagatcaac gagccacaag tactctcgga taatttagag 780  
gaattacgag ttgcattaaa cggtggcgac aaatataaac tagtctcacg gaagttagat 840  
gttgaatcta agaggaattt tgtgaaggaa tatatcagcg atcaacgtaa aaagaggaa 900  
tag 903

<210> 229  
<211> 2280  
<212> DNA  
<213> *Saccharomyces* sp.

<400> 229  
atgcctgcac caaaactcac ggagaaattt gcctcttcca agagcacaca gaaaactacg 60  
aattacagtt ccatcgaggc caaaagcgctc aagacgtcgg ctgatcaggc atacatctac 120

caagagccta gcgctacca gaagatactt tactccatcg ccacatggct gttgtacaac 180  
 atcttccact gcttctttag agaaatcaga ggccggggca gtttcaaggt accgcaacag 240  
 ggaccgggtga tctttgttgc ggctccgcat gctaaccagt tcgtcgaccc tgtaatcctt 300  
 atgggcgagg tgaagaaatc tgtaaacaga cgtgtgtcct tcttgattgc ggagagctca 360  
 ttaaagcaac ccccataggt gtttttggct agtttcttca tggccatagg cgtggtaagg 420  
 ccgcaggata atttgaaacc ggcagaaggt actatccgct tagatccaac agactacaag 480  
 agagttatcg gccacgacac gcatttcttg actgattgta tgccaaaggg tctcatcggg 540  
 ttacccaaat caatgggatt tggagaaatc cagtccatag aaagtgcacac gagtttgacc 600  
 ctaagaaaag agttcaaaat ggccaaacca gagattaaaa ctgctttact caccggcact 660  
 acttataaat atgccgctaa agtcgaccaa tcttgcgttt accatagagt ttttgagcat 720  
 ttggcccata acaactgcat tgggatcttt cctgaagggt ggtcccacga cagaacaaac 780  
 ttgttgcccc tgaaagcagg tgtggcgatt atggctcttg gttgcatgga taagcatcct 840  
 gacgtcaatg ttaagattgt tccctgcggt atgaattatt tccatccaca taagttcagg 900  
 tcgagagcgg ttgttgaatt cggtagcccc attgaaatac cgaaggaact agtcgccaaag 960  
 taccacaacc cggaaacgaa cagagatgca gtgaaagaat tattagatac catatcgaag  
 1020  
 ggtttacaat ccgttaccgt tacatgttct gattatgaaa ctttgatggg ggttcaaacg  
 1080  
 ataagaagac tatatatgac acaatttagc accaagttac cgttgccctt gattgtggaa  
 1140  
 atgaacagaa gaatgggtcaa aggttacgaa ttctatagaa acgatcctaa aatagcggac  
 1200  
 ttgaccaaag atataatggc atataatgcc gccttgagac actataatct tctgatcac  
 1260  
 cttgtggagg aggcaaagg aaatttcgca aaaaacctcg gacttgtttt ttttagatcc  
 1320  
 atcgggtctt gcattctctt ttcgttagcc atgccaggta tcattatgtt ctcacctgtc  
 1380  
 ttcatattag ccaagagaat ttctcaagaa aaggcccgtc ccgctttgtc caagtctaca  
 1440  
 gttaaaataa aggctaacga tgtcattgcc acgtggaaaa tcttgattgg gatgggattt  
 1500  
 gcgcccctgc tttacatctt ttgggtccgt ttaatcactt attacctcag acataaacca  
 1560  
 tggaaataaa tatatgtttt ttccgggtct tacatctcgt gtgttatagt cacgtattcc  
 1620  
 gccttaacgc tgggtgatat tggatggat ggtttcaa atcttgagacc actgggtttt  
 1680  
 tctcttacat ctccaaaggg cttgcaaaag ctacaaaagg atcgtagaaa tctggcgaaa  
 1740  
 agaataatcg aagttgtaa taactttgga agcgaattat tccccgattt cgatagtgcc  
 1800  
 gccctacgtg aagaattcga cgtcatcgat gaagaggaag aagatcgaaa aacctcagaa  
 1860  
 ttgaatcgca ggaaaatgct aagaaaacag aaaataaaaa gacaagaaaa agattcgtca  
 1920  
 tcacctatca tcagccaacg tgacaaccac gatgcctatg aacaccataa ccaagattcc  
 1980  
 gatggcgtct cattgggtcaa tagtgacaat tccctctcta acattccatt attctcttct  
 2040  
 acttttcac gtaagtcaga gtcttcctta gcttcgacat ccgttgacc ttcttcttcc  
 2100  
 tccgaatttg aggtagaaaa cgaaatcttg gaggaaaaa atggattagc aagtaaaatc  
 2160  
 gcacaggccg tcttaaacaa gagaattggg gaaaatactg ccaggaaga ggaagaggaa  
 2220  
 gaagaagagg aagaagaaga agaggaagaa gaagaagaag ggaaagaagg agatgcgtag  
 2280

&lt;210&gt; 230

&lt;211&gt; 2232

&lt;212&gt; DNA

<213> *Saccharomyces* sp.

&lt;400&gt; 230

atgtctgtc ccgtgcccga tcataacgct gccaaaccta ttctcatgt acctcaagcg 60  
 tcccgcgggt acaaaaattc atacaatgga ttcgatatac atatacatat atggctgtat 120  
 gatgtgtctg tatttctgtt taatatcttg ttactattt tcttcagaga aattaaggta 180  
 cgtggtgcat ataacgttcc cgaagtggg gtgccaacca tcttctgtgtg tgcccctcat 240  
 gcaaatcagt tcatcgaccc ggctttggta atgtcgcaaa cccgtttgct gaagacatca 300

```

gcgggaaagt cccgatccag aatgccttgt tttgttactg ctgagtcgag ttttaagaaa 360
agatttatct ctttctttgg tcacgcaatg ggcgggtattc ccgtgcctag aattcaggac 420
aacttgaagc cagtggatga gaatcttgag atttacgctc cggacttgaa gaaccacccg 480
gaaatcatca agggccgctc caagaaccca cagactacac cagtgaactt tacgaaaagg 540
ttttctgcca agtccttgct tggattgccc gactacttaa gtaatgctca aatcaaggaa 600
atccccgatg atgaaacgat aatcttgtcc tctccattca gaacatcgaa atcaaaagtg 660
gtggagctct tgactaatgg tactaatttt aaatatgcag agaaaatcga caatacggaa 720
actttccaga gtgtttttga tcaattgcat acgaagggct gtgtagggtat tttccccgag 780
ggtggtttctc atgaccgtcc ttcgttacta cccatcaagg caggtggttc cattatggct 840
ctgggcgagc tagccgctga tcctaccatg aaagtgtctg ttgtaccctg tggtttgcat 900
tatttccaca gaaataaatt cagatctaga gctgttttag aatacggcga acctatagtg 960
gtggatggga aatatggcga aatgtataag gactccccac gtgagaccgt ttccaaacta
1020
ctaaaaaaga tcaccaattc tttgttttct gttaccgaaa atgctccaga ttacgatact
1080
ttgatgggtca ttcaggtctg cagaagacta tatcaaccgg taaaagtcag gctacctttg
1140
cctgccattg tagaaatcaa cagaagggtta cttttcgggtt attccaagtt taaagatgat
1200
ccaagaatta ttcacttaaa aaaactggta tatgactaca acaggaaatt agattcagtg
1260
ggtttaaaag accatcaggt gatgcaatta aaaactacca aattagaagc attgaggtgc
1320
tttgtaactt tgatcgttcg attgattaaa ttttctgtct ttgctatact atcggttaccg
1380
ggttctatct tcttcaactc aattttcatt atttgcgcg tatactcaga aaagaaggcc
1440
aaagagggtt taaagaaatc attgggttaa attaagggtta ccgatttggt ggccacatgg
1500
aaacttatcg tggcgtaat attggcacca attttatacg ttacttactc gatcttgggtg
1560
attattttgg caagaaaaca acactattgt cgcactctggg ttccttccaa taacgcattc
1620
atacaatttg tctattttta tgcgttattg gttttcacca cgtattcctc tttaaagacc
1680
ggtgaaatcg gtgttgacct tttcaaactc ttaagaccac tttttgttcc tattgtttac
1740
cccggtgaag agatcgaaga aatccaaaca acaagaaaga atttaagtct agagttgact
1800
gctgtttgta acgatttagg acctttgggt ttccttgatt acgataaatt agcgactgag
1860
atattctcta agagagacgg ttatgatgtc tcttctgatg cagagtcttc tataagtcgt
1920
atgagtgtac aatctagaag ccgctcttct tctatacatt ctattggctc gctagcttct
1980
aacgccctat caagagtga ttcaagagggc tcgttgaccg atattccaat tttttctgat
2040
gcaaagcaag gtcaatggaa aagtgaaggt gaaactagtg aggatgagga tgaatttgat
2100
gagaaaaatc ctgccatagt acaaaccgca cgaagtctct atctaaataa ggaaaaacagt
2160
cgcaacacaa atatatcttc gaagattgct tcgctggtaa gacagaaaag agaacacgaa
2220
aagaaagaat ga
2232

```

&lt;210&gt; 231

&lt;211&gt; 1194

&lt;212&gt; DNA

<213> *Saccharomyces* sp.

&lt;400&gt; 231

```

atgctgcacg aaaaaatagc tcataaagtt cgaaaagtcg tcgtcccagg tatttcctta 60
ttgattttct tccagggatg ctttatttct ttgtttctcc aactcaccta taagactctt 120
tactgtagaa atgatataag gaaacaaatt ggtctcaata aaaccaaagc attatttatt 180
gtcttggtat catccatttt gcatgttgct gcaccatctg cagtgagaat taccactgaa 240
aattccagtg ttcctaaagg tacttttttt ttagacttga agaagaaaag gattctttct 300
catctaaagt ccaattcggt ggccatttgc aatcaccaaa tatacacgga ttggatattt 360
ttatggtggt tggcttacac atcgaaacta ggggctaagt tcttcattat tttaaaaaaa 420
tcgttggctt ccattcctat cctcgggttc ggtatgagaa actataattt catttttatg 480

```



agtagaaagt gggcacaaga caaaataac ctaagcaaca gccttgctgg ccttgattcg 540  
aatgcaaggg gcgccggctc acttgctgga aagtcacctg agcgcataac tgaggaagga 600  
gagagcatat ggaatccgga ggttattgat ccaaaacaaa tccattggcc atacaatctt 660  
atcctattcc ctgaagggtac aaatctcagt gctgatacta ggcaaaaaag tgctaaatat 720  
gctgccaaaa taggcaaaaa gccattcaag aatgtgctac tgcctcattc tacaggccta 780  
agatactcgt taaaaaagtt gaagccaagt attgaaagtc tttatgatat tacgatcggc 840  
tactccgggtg taaaacagga ggaatatggt gagcttatat atgggctgaa gagcatattt 900  
ttagaaggaa aatacccgaa gttagtcgat attcacatca gagcatttga tgtaaagat 960  
attccattag aggacgagaa tgaattttca gaatggctgt ataaaatttg gagtgagaag  
1020  
gatgctctaa tggaaaggta ctattccact ggatcattcg taagtgatcc tgaacaaaac  
1080  
cattcagtta ccgatagttt caagatcaat cgtattgagt taactgaagt gctaataatta  
1140  
ccaactctaa caataatttg gttagtttat aaactttatt gttttatttt ttga  
1194

&lt;210&gt; 232

&lt;211&gt; 912

&lt;212&gt; DNA

&lt;213&gt; Saccharomyces sp.

&lt;400&gt; 232

atgagtgtga taggtaggtt cttgtattac ttgaggtccg tgttggtcgt actggcgctt 60  
gcaggctgtg gcttttacgg tgtaatcgcc tctatccttt gcacgttaat cggtaagcaa 120  
catttggtctc agtggattac tgcgcgttgt ttttaccatg tcatgaaatt gatgcttggc 180  
cttgacgtca aggtcgttgg cgaggagaat ttggccaaga agccatataat tatgattgcc 240  
aatcaccaat ccaccttggga tatcttcatg ttaggttagga ttttcccccc tggttgcaca 300  
gttactgccca agaagtcttt gaaatacgtc ccctttctgg gttggttcat ggctttgagt 360  
ggtacatatt tcttagacag atctaaaagg caagaagcca ttgacacctt gaataaagggt 420  
ttagaaaatg ttaagaaaaa caagcgtgct ctatgggttt ttcttgaggg taccagggtc 480  
tacacgagt agctgacaat gttgcctttc aagaagggtg ctttccattt ggcaaacag 540  
ggtaagatcc ccattgttcc agtggttgtt tccaatacca gtactttagt aagtcctaaa 600  
tatgggggtct tcaacagagg ctgtatgatt gttagaattt taaaacctat ttcaaccgag 660  
aacttaacaa aggacaaaat tggatgaatt gctgaaaaag ttagagatca aatgggtgac 720  
actttgaagg agattggcta ctctccgcc atcaacgata caaccctccc accacaagct 780  
attgagtatg ccgctcttca acatgacaag aaagtgaaca agaaaatcaa gaatgagcct 840  
gtgccttctg tcagcattag caacgatgtc aatacccata acgaagggtc atctgtaaaa 900  
aagatgcatt aa 912

&lt;210&gt; 233

&lt;211&gt; 54

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence:Synthetic  
Oligonucleotide

&lt;400&gt; 233

cgcgatttaa atggcgcgcc ctgcaggcgg ccgcctgcag ggcgcgccat ttaa 54

&lt;210&gt; 234

&lt;211&gt; 32

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence:Synthetic  
Oligonucleotide

&lt;400&gt; 234

tcgaggatcc gcggccgcaa gcttcctgca gg 32

&lt;210&gt; 235

&lt;211&gt; 32

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence:Synthetic  
Oligonucleotide

<400> 235  
tcgacctgca ggaagcttgc ggccgcggat cc 32

<210> 236  
<211> 32  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:Synthetic  
Oligonucleotide

<400> 236  
tcgacctgca ggaagcttgc ggccgcggat cc 32

<210> 237  
<211> 32  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:Synthetic  
Oligonucleotide

<400> 237  
tcgaggatcc gcggccgcaa gcttcctgca gg 32

<210> 238  
<211> 36  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:Synthetic  
Oligonucleotide

<400> 238  
tcgaggatcc gcggccgcaa gcttcctgca ggagct 36

<210> 239  
<211> 28  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:Synthetic  
Oligonucleotide

<400> 239  
cctgcaggaa gcttgccggcc gcggatcc 28

<210> 240  
<211> 36  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:Synthetic  
Oligonucleotide

<400> 240  
tcgacctgca ggaagcttgc ggccgcggat ccagct 36

<210> 241  
<211> 28  
<212> DNA  
<213> Artificial Sequence

&lt;220&gt;

<223> Description of Artificial Sequence: Synthetic  
Oligonucleotide

&lt;400&gt; 241

ggatccgcgg ccgcaagctt cctgcagg

28